Zonation of ribosomal DNA transcription defines a stem cell hierarchy in colorectal cancer

Clara Morral[1,9], Jelena Stanisavljevic[1,9], Xavier Hernando-Momblona[1,8], Elisabetta Mereu[2,] , Adrián Álvarez[1,8] , Carme Cortina[1,8], Diana Stork[1] , Felipe Slebe[1] , Gemma Turon[1], Gavin Whissell[1], Marta Sevillano[1,8], Anna Merlos-Suárez[1], Angela Casanova-Martí[1], Catia Moutinho[2], Scott W. Lowe[4], Lukas E. Dow[5], Alberto Villanueva[6], Elena Sancho[1,8], Holger Heyn[2,3], Eduard Batlle[1,7,8,10]

1. Institute for Research in Biomedicine (IRB Barcelona). The Barcelona Institute of Science and Technology (BIST), Baldiri i Reixac 10, 08028 Barcelona, Spain.
2. CNAG-CRG, Centre for Genomic Regulation (CRG), Barcelona Institute of Science and Technology (BIST), Baldiri i Reixac 4, 08028 Barcelona, Spain.
3. Universitat Pompeu Fabra (UPF), Barcelona, Spain.
4. Department of Cancer Biology and Genetics, Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA; Howard Hughes Medical Institute, Chevy Chase, MD 20815, USA.
5. Department of Medicine, Weill-Cornell Medical College, New York, NY 10021, USA.
6. Group of Chemoresistance and Predictive Factors, Subprogram Against Cancer Therapeutic Resistance (ProCURE), ICO, Oncobell Program, IDIBELL, L'Hospitalet del Llobregat, 08908 Barcelona, Spain.
7. ICREA, Pg. Lluís Companys 23, 08010 Barcelona, Spain.
8. Centro de Investigación Biomédica en Red de Cáncer (CIBERONC), Barcelona, Spain.
9. These authors contributed equally.
10. Lead Contact

Lead contact: eduard.batlle@irbbarcelona.org

SUMMARY

Colorectal cancers (CRCs) are composed of an amalgam of cells with distinct genotypes and phenotypes. Here we reveal a previously unappreciated heterogeneity in the biosynthetic capacities of CRC cells. We discover that the majority of ribosomal DNA transcription and protein synthesis in CRCs occur in a limited subset of tumor cells that localize in defined niches. The rest of the tumor cells undergo an irreversible loss of their biosynthetic capacities as a consequence of differentiation. Cancer cells within the biosynthetic domains are characterized by elevated levels of the RNA Polymerase 1 subunit A - POLR1A. Genetic ablation of POLR1A-high cell population imposes an irreversible growth arrest to CRCs. We show that elevated biosynthesis defines stemness in both LGR5+ and LGR5- tumor cells. Therefore, a common architecture in CRC is a simple cell hierarchy based on the differential capacity to transcribe ribosomal DNA and synthesize proteins.

INTRODUCTION

Only subsets of tumor cells isolated from CRC samples, the so-called cancer stem cells (CSCs), display tumorigenic potential upon transplantation into recipient mice (Dalerba et al., 2007; Merlos-Suárez et al., 2011; O'Brien et al., 2007; Ricci-Vitiani et al., 2007; Vermeulen et al., 2010). This finding lead to the notion that CRCs are organized according to a stem cell hierarchy. Indeed, tumor cells present in full-blown CRCs echo the phenotypes observed in healthy colonic mucosa, including stem cell-like and differentiated-like states characterized by expression programs reminiscent of those present in their normal counterparts (Dalerba et al., 2011; Merlos-Suárez et al., 2011; Vermeulen et al., 2008). Lineage-tracing experiments in xenografts demonstrated that LGR5+ stem cell-like tumor cells exhibit long-term self-renewal and differentiation capacities whereas KRT20+ differentiated cells produce mostly short-lived progeny (Cortina et al., 2017; Shimokawa et al., 2017). However, genetic ablation of LGR5+ cells in these cancer models triggers a regenerative response by which LGR5-negative tumor cells convert to an LGR5+ state and regenerate the CSC pool (Shimokawa et al., 2017; de Sousa e Melo et al., 2017). In addition, de Sauvage and colleagues showed that ablation of LGR5+ cells does not impair primary CRC growth yet these cells are necessary for the expansion of metastases (de Sousa e Melo et al., 2017). These studies raise a number of important caveats. In particular, CSCs have been recognized by expression of stem cell marker gene LGR5 yet many CRCs contain few or no LGR5+ cells (Merlos-Suárez et al., 2011; Shimokawa et al., 2017). It is unclear whether these LGR5-negative CRCs still rely on a CSC hierarchy or not. One possibility is that CSCs may not be simply defined by LGR5 expression in some CRCs or even that particular tumors may contain LGR5+ and LGR5-negative CSC subsets. Furthermore, although evidence supporting tumor cell plasticity is well established, it remains unknown whether all cell types present in CRCs are equally capable of regenerating the CSCs pool. We reasoned that understanding the functions that characterize CSCs beyond expression of marker genes might help address these important issues.

While characterizing the properties of CSCs, we found that most rRNA and proteins synthesized in CRCs are contributed by a limited subset of cells that reside immediately adjacent to the stroma. In contrast, as tumor cells undergo differentiation, they experience an irreversible loss of rRNA and protein synthesis capacity. By exploring these unexpected findings, we demonstrate that zonation patterns of rDNA transcription and protein synthesis in CRC reflect the existence of a

simple stem cell-like hierarchy based on differential biosynthetic capacity of tumor cells.

RESULTS

Zonation of rRNA and protein synthesis in CRC

We discovered striking zonation patterns of protein and ribosomal RNA (rRNA) synthesis in CRCs. O-propargyl-puromycin (OP-P) is a puromycin analog that is incorporated into nascent polypeptides (Blanco et al., 2016; Liu et al., 2012; Signer et al., 2014). In CRC patient-derived xenografts (PDX), the OP-P+ domain was visualized as strings of cells adjacent to the stroma that surrounds each tumor gland (Figure 1A). Closer inspection confirmed that OP-P incorporation was largely restricted to tumor cells occupying basal positions, with the highest protein synthesis rates corresponding to those cells in closest proximity to the underlying mesenchyme whereas cells that reside outside these domains, including KRT20+ differentiated tumor cells, did not incorporate OP-P (Figure 1A' and A''). Stromal cells stained strongly for OP-P. These OP-P zonation patterns were present in all xenografts analyzed that included subcutaneous, orthotopic and metastatic CRC models (Examples in Figure 1B, 1C, S1A, and S1B). We also measured rDNA transcription by inoculating mice with 5-Ethynyl Uridine (EU), a ribonucleotide analog that can be visualized using click chemistry (Jao and Salic, 2008). EU incorporates mainly into rRNA (Jao and Salic, 2008) and, similar to OP-P, it labeled preferentially the nucleoli of tumor cells located adjacent to tumor stroma after a 2 h pulse (Figure 1D, 1D' and 1D'', S1C and S1D).

We next searched for surface marker genes that overlapped with CRC biosynthetic zones identified above. In CRCs, highest EPHB2 levels are expressed by a subset of undifferentiated tumor cells that reside in proximity to the stroma (Batlle et al., 2002; Merlos-Suárez et al., 2011). EPHB2-high cells resided within OP-P+ tumor domains (Figure S1E and S1F). Flow cytometry of dissociated xenografts demonstrated that EU and OP-P incorporation was largely restricted to EPHB2-high tumor cells and further confirmed that most EPHB2-low tumor cells exhibited near-zero rDNA transcription and protein synthesis rates (Figure 1E and 1F). We obtained similar measurements in three xenografts from tumors with distinct genotypes (Figure 1G). We also performed transcriptional profiling of tumor cells purified from fresh surgical CRC specimens according to EPHB2 surface levels (n=10) (Figure 1H). This study confirmed significant enrichment in gene sets associated to rRNA and protein

synthesis in EPHB2-high compared to -low tumor cells in all samples (Figure 1I and Table S6). Overall, these findings imply the existence of well-defined biosynthetic domains in CRC dedicated to producing rRNA and proteins (Figure 1J). This organization occurs independently of the site of growth and it is a defining feature of many CRCs.

Characterization of biosynthetic cells in CRC

Recent reports revealed that CRCs expand through the proliferation of tumor cells that are located at the tumor edges in contact with the stroma that surrounds xenografts (van der Heijden et al., 2019; Lenos et al., 2018). In our model systems, however, there were no significant differences in the numbers of Ki67+ cells located at the tumor center and in the periphery (Figure S2A, S2B, S2C and S2D). We confirmed this finding in CRC patient samples (Figure S2E and S2F). The OP-P+ and EU+ domains only encompassed a fraction of the Ki67+ tumor cell population and were similarly distributed in the tumor center and the periphery (Figure 2A and 2B).

Previous studies have shown that a substantial proportion of CRCs exhibited few or no LGR5+ cells (Figure 2C and 2D) (Merlos-Suárez et al., 2011; Shimokawa et al., 2017). For convenience, we reproduce in Figure 2E the results from Merlos-Suárez et al. We investigated how the expression of LGR5 correlated with rRNA and protein synthesis zonation patterns described above. To this end, we knocked-in a tdTomato cassette in the LGR5 locus of patient-derived organoids (PDOs) using a CRISPR/Cas9-based approach that we described elsewhere (Cortina et al., 2017) (Figure 2F). In these experiments, we used two different PDOs as model systems: PDO-p18 carries inactivating mutations in APC and functional inactivation of P53 (van de Wetering et al., 2015) and contains a large proportion of LGR5+ cells. PDO7 is mutant in four main driver pathways (APC, KRAS, ATM, SMAD4) and represents a model for CRCs with a discrete LGR5+ cell population (Table S4). Subsequently, we generated xenografts inoculating these LGR5 reporter PDOs in immunodeficient mice. tdTomato-high cells purified from dissociated xenografts upregulated ISC-specific genes such as LGR5 and SMOC2 whereas tdTomato-negative cells expressed markers of intestinal differentiation (Figure 2G and 2J). PDO-p18-derived xenografts contained abundant LGR5+ cells (Figure 2H) and we observed overlap between the LGR5+ and OP-P+ domains in these tumors (Figure 2I). In contrast, PDO7 xenografts exhibited a smaller proportion of LGR5+ cells scattered through

tumor glands (Figure 2K), most of which were distributed within the OP-P-low/negative domain (Figure 2L).

It was also evident that tumor glands formed by differentiated KRT20+ tumor cells did not incorporate OP-P or EU, even if these differentiated cells were immediately adjacent to stroma (examples in Figure 1A-D and 2M). The lumens of these KRT20+ glands often contained dead cells (Figure 2N and 2N') suggesting that differentiated CRC cells exhibit short lifespans as previously proposed (Shimokawa et al., 2017). In addition, most CRCs analyzed included a small proportion of KRT20+ (0.5% - 2%) scattered within the OP-P+ domain (Figure 2O, 2O' and 2O''). We conclude that the EU+ and OP-P+ domains of CRCs are always positioned adjacent to the stroma yet the stromal localization *per se* neither directly defines the biosynthetic capacities of CRC cells, nor the expression of stem cell, differentiation or proliferation marker genes.

Shutdown of rRNA and protein synthesis rates during tumor cell differentiation

Tumor cell differentiation can be enforced in CRC cell lines and organoids through the inhibition of WNT signaling (van de Wetering et al., 2002). Expression of a tamoxifen-inducible dominant negative TCF4 transcription factor in LS174T or SW403 CRC cells triggered the expression of KRT20 and other differentiation markers coinciding with a decline in the levels of ISC genes (Figure 3A). Likewise, mouse intestinal tumoroids engineered to bear activating mutations in Kras (G12D) combined with loss of function alleles of p53 and a doxycycline-regulated shRNA targeting Apc (AKP tumoroids) (Dow et al., 2015) undergo differentiation upon removal of doxycycline leading to shutdown of beta-catenin/TCF4 transcription (Figure 3B). We observed that after prolonged WNT blockade and differentiation (7 days), CRC cell lines and mouse tumoroids exhibited a significant reduction in EU and OP-P incorporation (Figure 3C-3D, S3A and S3B). The uptake of methionine analog L-homopropargylglycine (HPG) was also largely reduced (Figure 3E and 3F and below). In addition, prolonged WNT blockade caused a drop in pre-rRNA levels (Figure 3G).

The Y10B antibody recognizes mainly cytoplasmic 5.8S rRNA loaded in ribosomes (Lerner et al., 1981; Witte et al., 1991). Using this antibody, we found that loss of OP-P and EU incorporation coincided with a decrease in the number of assembled ribosomes (Figure 3H). Y10B staining of primary formalin fixed-paraffin embedded (FFPE) patient samples (n=24) revealed extensive intratumoral heterogeneity in

ribosome density. Elevated expression of the pan-differentiation marker KRT20 correlated strongly with lack of 5.8S staining in the vast majority of patients (Example in Figure 3I-3J, and S3C-S3D). Hence, KRT20+ CRC cells not only display a pronounced reduction in rRNA synthesis and protein synthesis rates but also exhibit a reduced number of ribosomes.

To track production of proteins during cell differentiation, we made use of HPG, which is incorporated into proteins without disturbing their functions (Calve et al., 2016; Dieterich et al., 2010; Signer et al., 2014). CRC cell lines expressed KRT20 protein 4 days after blockade of the WNT pathway (Figure 3L). At this early time point, many KRT20+ cells incorporated HPG implying that during the onset of differentiation protein synthesis rates remain elevated. Yet, we observed a progressive decline in HPG incorporation during the following 3 days that coincided with downregulation of rRNA and protein synthesis machinery (Figure 3K and 3L). Pulse-chase experiments revealed that the majority of KRT20 was synthesized during the first three days after WNT blockade yet only limited amounts were produced at a later time point (Figure 3M-3P). Therefore, fully differentiated tumor cells retain long-lived proteins that were synthesized before the protein synthesis machinery was shut down.

Biosynthetic CRC cells express elevated POLR1A levels

We noticed that prolonged WNT blockade in *in vitro* CRC models caused downregulation of mRNAs encoding for subunits of RNA Polymerase-I holocomplex including POLR1A, POLR1B and POLR1C (Figure 4A and 4B). The rDNA specific transcription factors TIF-1A and UBTF were also downregulated (Figure 4B). We confirmed substantial reduction of the major subunit of the RNA Polymerase 1 complex - POLR1A - after 7 days of WNT blockade by western-blot (Figure 4C) and by immunofluorescence (Fig 4D and S4A). Consistently, KRT20+ cells present in xenografts displayed relatively low POLR1A levels (Figure S4B and S4C). The expression of POLR1A was increased in CRC samples compared to normal mucosa, which is consistent with augmented rDNA transcription in tumors (Figure S4D).

To study POLR1A *in vivo*, we used CRISPR/Cas9 technology to knock-in an *EGFP* cassette in frame with the start codon of *POLR1A* gene (Figure 4E). The EGFP-POLR1A fusion protein was located in the nucleolus of knock-in organoids (Figure 4F and S4E). Upon inoculation into immune-deficient mice, PDOs generated xenografts displaying extensive EGFP-POLR1A heterogeneity. EGFP-POLR1A and KRT20+

areas were located in mutually exclusive domains in PDXs (Figure 4G and S4F). Isolation of the EGFP-POLR1A-high cell population (brightest 10% tumor cells) from dissociated tumors followed by western blot confirmed elevated POLR1A levels whereas their negative counterparts expressed KRT20 (Figure 4H, 4I, S4G and S4H). EU and OP-P were almost exclusively incorporated in EGFP-POLR1A-high cells purified from both PDO7 and PDO-p18 xenografts (Figure 4J-4N and S4I). Therefore, elevated POLR1A levels characterize the cells that reside within the biosynthetic domains of CRC.

EGFP-POLR1A-high cells purified from xenografts exhibited higher organoid forming capacity than EGFP-POLR1A-low cells (Figure S4J). Importantly, experiments of tumor cell isolation followed by transplantation into secondary recipients revealed that the EGFP-POLR1A-high cell population present in both PDO-p18 and PDO7 was largely enriched in tumor-initiating cells (Figure 4O). Intrasplenic inoculation of tumor cell populations also demonstrated that the vast majority of metastasis initiating cells reside within the tdTomato-POLR1A-high cell fraction (Figure S4K and S4L).

Characterization of tumor cell heterogeneity by single-cell profiling

We next studied the relationship between stem and differentiated gene programs and the POLR1A-high cell population by single-cell RNA sequencing (scRNA-seq) of PDO7- and PDO-p18-derived xenografts. Unsupervised clustering identified discrete tumor cell populations that we subsequently ascribed to either ISC-like or differentiated-like (absorptive, enteroendocrine and mucosecreting) phenotypes (Figure 5A, 5F, S5A and S5B). The LGR5 expression signature labeled ISC-like cells whereas KRT20 was upregulated in differentiated-like tumor cells (Figure 5B, 5G, S5A and S5B). These analyses also identified a distinctive population of undifferentiated cells that expressed elevated POLR1A-high signature levels in both xenografts (Figure 5C and 5H). The transcriptomes of the POLR1A-high cell population in both PDO-P18 and PDO7 xenografts were enriched in gene sets related to proliferation and biosynthetic pathways (Table S7). Crucially, the two PDO xenografts differed in the expression patterns of the LGR5 signature. Many tumor cells in PDO-p18, including a large fraction of POLR1A-high cell population, were ISC-like and expressed LGR5 (Figure 5B and 5C). In contrast, there was a limited overlap between the POLR1A-high and the LGR5+ cell populations in PDO7 (Figure 5G and 5H). Using PDO7 bearing LGR5-tdTomato and POLR1A-EGFP reporter cassettes (Figure S4M) we corroborated that only a minority (10%) of POLR1A-

EGFP-high cells were LGR5-tdTomato+ cells (Figure S4N). Within the POLR1A-high population, LGR5+ and LGR5- cells exhibited differential expression of metabolism-related gene sets (Figure S4O and S4P).

Finally, analysis of gene expression dynamics using pseudotime reconstruction predicted that the origin of the trajectory in both PDO7 and PDO-p18 xenografts corresponded to the POLR1A-high cells whereas differentiated-like cells were positioned at endpoints (Figure 5D, 5E, 5I and 5J). Remarkably, changes in cell states over the pseudotime path were characterized by a progressive decline of POLR1A-high signature expression levels (Figure S5C-F).

Clonal analysis of POLR1A+, LGR5+ and KRT20+ cells

By means of CRISPR/Cas9, we knock-in lineage tracing cassettes into endogenous *POLR1A*, *LGR5* or *KRT20* locus of PDO7 (Figure 5K). We titrated the amount of tamoxifen necessary to recombine approximately the same number of cells (1-2%) in the three creERT2 PDO-derived xenografts (Figure S5G). This low recombination frequency facilitated the analysis the three cell types at the clonal level (Figure 5P). We found no significant differences in the fraction of non-necrotic tumor cells between the three experimental conditions (Figure S5H). The number of clones generated by POLR1A+ cells was sustained over time whereas the frequency LGR5-derived clones declined at late time points (Figure 5L). In addition, POLR1A+ tumor cells produced larger clones (Figure 5M, 5N and 5P) and contributed substantially more cells than LGR5+ cells (Figure 5O). Our data also indicates that both POLR1A and LGR5 populations generated progeny that underwent differentiation as shown by KRT20 labeling (Figure 5Q). Based on these observations, we tentatively conclude that POLR1A+ biosynthetic cells represent the main population fueling the growth of PDO7 tumors.

Most KRT20+ differentiated cells exhibited low proliferative potential and were short-lived (Figure 5L, 5M and 5O). Yet, a small proportion of KRT20+ cells (2-5%) generated progeny that expanded over time (Figure 5N), which may indicate that either some cells retained proliferative potential or that a restricted subset of differentiated cells recovered stem cell potential thorough plasticity. It is, however, important to consider that at the earliest time point of tamoxifen treatment (3 days), only a minority (15%) of the tdTomato marked cells represented mature differentiated cells as shown by lack of expression of KRT20 protein (Figure S5I and S5J). We

speculate that due to the drop of protein synthesis rates, the probability of inducing recombination in terminally differentiated cells is low.

POLR1A-high cells are necessary for tumor growth

We next performed experiments of cell ablation using PDOs that carry an inducible caspase 9 (iCasp9) knocked-in in either the *POLR1A* or the *LGR5* locus (Shimokawa et al., 2017) (Figure 6A, 6B, 6J and 6N). Xenografts derived from inoculation of knock-in PDOs in mice exhibited heterogeneous POLR1A-tdTomato-iCasp9 (POLiCT) levels, including KRT20+ cells that express very low levels of the reporter (Figure 6C and 6D). Indeed, analysis of tdTomato by flow cytometry confirmed a wide range of expression, with bright and dim cells differing more than 25-fold in reporter levels (Figure 6E). POLR1A-tdTomato-high cells expressed higher levels of POLR1A (Figure S6A) and exhibited increased OP-P and EU incorporation compared to POLR1A-tdTomato-low cells (Figure S6C and S6D). Induction of iCasp9 dimerization with AP20187 (Shimokawa et al., 2017) selectively eliminated the brightest POLR1A-tdTomato cells after 5 days (Figure 6E). In addition, we observed an increased frequency of cells expressing low reporter levels suggesting tumor cell differentiation. By immunofluorescence, we confirmed that only tumor cells expressing the highest Tomato levels were ablated in these experiments (Figure S6B). Upon treatment, the epithelial compartment, but not the stromal cells of these xenografts lacked OP-P and EU incorporation (Figure 6G, S6D) demonstrating that we effectively eliminated the cell population that synthesized proteins and rRNA in tumors. Ablation of POLR1A-high cells completely halted the growth of both PDO-p18 and PDO7 xenografts (Figure 6F). It also caused an increase in the number of differentiated cells shown by KRT20 and Alcian blue staining, although this effect was more prominent in PDO7 than in PDO-p18 (Figure 6H, 6I, S6E and S6F). Genetic ablation of LGR5 cells (Figure 6K, 6O, S6G and S6H) using an equivalent approach was effective, but only halted growth and OP-P incorporation in PDO-p18 xenografts (Figure 6L and 6M). PDO7 tumors grew steadily and the zonation pattern of protein synthesis was not altered despite lack of LGR5+ cells (Figure 6P and 6Q). Overall, these results imply that biosynthetic cells labeled by high POLR1A levels sustain tumor growth in both CRCs. The broad expression of LGR5 in PDO-p18, which expands into the POLR1A+ compartment, makes this tumor susceptible to LGR5+ cell ablation. In contrast, only a minority of LGR5+ cells in PDO7 express POLR1A and exhibit biosynthetic capacities and, therefore, the bulk of the LGR5+ cell population is dispensable for tumor growth.

Therapeutic targeting of biosynthetic tumor cells

Treatment of xenografts with FOLFIRI, a chemotherapeutic commonly used in CRC patients, slowed down tumor growth and reduced the number of POLR1A-tdTomato-high cells (Figure S6I, S6J and S6K). Incorporation of OP-P+ was also diminished (Figure S6L). Concomitantly, FOLFIRI-treated xenografts exhibited an increased abundance of POLR1A-tdTomato-Low cells (Figure S6K) and contained substantially more KRT20+ cells at experimental endpoints (Figure S6L and S6M). Therefore, FOLFIRI treatment reproduces to some extent the effects of genetic ablation of POLR1A-high cells although it failed to completely eliminate the tumor biosynthetic compartments.

BMH-21 is a small molecule that induces degradation of POLR1A (Peltonen et al., 2014). *In vitro* treatment with BMH21 downregulated POLR1A levels and induced overt differentiation of PDO7 (Figure S6N, S6O and S6P). We observed similar yet milder effects in xenografts (S6Q and S6R). Higher BMH21 doses resulted in mouse toxicity, which limited its therapeutic efficacy in the *in vivo* setting.

Irreversible phenotype of KRT20+ CRC cells

Experiments of POLR1A-high cell ablation during 10 days followed by a long recovery period in the absence of dimerizer revealed that growth remained continuously halted (Figure 7A and 7B). Analysis of the remaining tumor mass 7 days after stopping dimerizer treatment showed the presence of KRT20+ tumor cells in direct contact with the stroma that exhibited largely reduced protein synthesis rates (Figure 7C). After extended periods of time (>25-30 days), tumors slowly reinitiated growth, which was likely due to residual cells that acquired resistance to iCasp9 ablation as inferred by the fact that tumors continuously treated with dimerizer expanded with almost identical kinetics (Data not shown) and that organoids established from samples taken at the onset of re-growth phase were resistant to dimerizer treatment *in vitro* (Figure S7A and S7B).

To explore further the irreversibility of the differentiated state, we made use of AKP tumoroids bearing a doxycycline-regulated *shApc* cassette (Figure 7D). Tumoroids forced to differentiate *in vitro* during 2 days in the absence of doxycycline (APC ON) that were subsequently cultured for 7 days with doxycycline (APC OFF), partially retained rRNA transcription, protein synthesis and growth rates (Figure 7E-7G). Therefore, loss of biosynthetic capacity due to short-term differentiation (2d) is reversible to some extent. In contrast, switching to doxycycline media after 5 days of

differentiation did not restore growth (Figure 7G and S7D), despite re-expression of the GFP-*Apc* shRNA cassette (Figure S7C). Most cells in these tumoroids were KRT20+ (Figure 7H) and exhibited reduced capacity to synthesize rRNA or proteins even after prolonged culturing in doxycycline-supplemented media (i.e. 7 days) (Figure 7E and 7F). Consistent with this finding, AKP tumoroids did not recover POLR1A levels after long-term (7d) differentiation (Figure 7H and 7I). Transcriptomic profiling of mouse tumoroids at early (2 days of Apc restoration) versus late stage differentiation (7 days of Apc restoration) revealed that downregulation of genesets related to biosynthesis such as ribosome, nucleolus, ribonucleoprotein assembly, ribosome biogenesis, etc., occurred only 7 days after restoring APC, concordantly with the time frame of irreversible differentiation (Figure S7F and S7G).

To study reversibility *in vivo*, we generated xenografts by inoculation of AKP tumoroids in mice (Figure 7J). Removal of doxycycline halted tumor growth, upregulated expression of KRT20 and reduced EU and HPG incorporation in xenografts (Figure 7K, 7L and S7E). This phenomenon was partially reversible by doxycycline supplementation after 8 days, including restoration of HPG and EU incorporation. Yet, 20 days after continuous APC expression, the addition of doxycycline did not rescue tumor growth, rRNA or protein syntheses rates (Figure 7K, 7L and S7E). The remaining tumor mass was composed mostly of KRT20+ cells directly in contact with the tumor stroma. POLR1A levels remained downregulated at this later time point (Figure 7M). Therefore, prolonged differentiation causes permanent downregulation of the biosynthetic capacities of CRC cells (Figure 7N).

DISCUSSION

Homeostasis of adult tissues requires controlled protein synthesis rates and distinct cell types in healthy tissues, including stem cells and their differentiated progeny, exhibit different biosynthetic capabilities (Blanco et al., 2016; Signer et al., 2014). In tumors, this regulation is disrupted by oncogenic alterations, many of which enhance the cell biosynthetic machinery, including rDNA transcription, ribosomal biogenesis, and protein production rates (Bustelo and Dosil, 2018; Pelletier et al., 2017; Ruggero, 2012). Genetic manipulations leading to diminished ribosomal function and protein production exert robust anti-tumorigenic effects (Ajore et al., 2017; Barna et al., 2008; Peltonen et al., 2014; Signer et al., 2014). Notwithstanding the high biosynthetic requirements of cancers, our data evidence a progressive loss of biosynthetic activity due to the pervasive differentiation that tumor cells experience in CRCs. Fittingly, in models of WNT-depend pancreatic cancer, inhibition of WNT signaling using Porcupine inhibitors causes differentiation (Jiang et al., 2013), and this process is accompanied by downregulation of POLR1A, POLR1B and other genes involved in ribosomal biogenesis (Madan et al., 2018). It is well established that signaling pathways that promote biosynthesis such as mTOR enhance cap-dependent translation, whereas stress conditions, including nutrient deprivation or hypoxia can activate alternative modes of translation, e.g., via internal ribosome entry sites (Robichaud and Sonenberg, 2017; Robichaud et al., 2019). Thus, the vastly reduced protein synthesis rates present in terminally differentiated cells may still be sufficient to translate specific mRNAs. Likewise, high protein synthesis rates shown by POLR1A-high cells may favor the translation of RNA subsets important for the biology of cancer stem cells.

Many CRCs contain abundant LGR5+ cells, yet lineage tracing analysis suggests that only a small proportion of them function as CSCs (Kozar et al., 2013). Our results support this conclusion. Other CRCs contain few to no LGR5+ cells (Merlos-Suárez et al., 2011; Shimokawa et al., 2017) but may still exhibit a hierarchical organization. The same caveats could be possibly applied to other putative stem cell marker genes. In addition, LGR5- cells present in primary CRCs act as seeds for metastasis in experimental models (Fumagalli et al., 2020). Our study defines functionally the CSC population in CRC: we propose that CSCs represent a subset of tumor cells with elevated rDNA transcription rates and protein synthesis capacity dedicated to fuel tumor growth. We show that both LGR5+ and LGR5- tumor cells that display elevated biosynthetic features function as CSCs in transplantation

experiments and also in intact tumors. The finding that the biosynthetic cell population occupies a basal position within the tumor glands suggests a role for the underlying mesenchymal cells in instructing their phenotype. However, proximity to this niche does not define unequivocally an elevated biosynthetic capacity as many CRCs contain KRT20+ and LGR5+ cells in direct contact with the stroma that nevertheless exhibit low protein and rRNA synthesis rates. Our findings fit well with a model whereby the properties of cancer cells are defined both by the microenvironment as well as by cell-intrinsic programs dictated by a stem cell hierarchy.

Previous experiments of KRT20+ tumor cell fate mapping have shown that this cell population gives rise mostly to short-lived progeny but can revert to an LGR5+ state upon ablation of this CSC pool (Shimokawa et al., 2017). It is however important to consider that at the onset of differentiation CRC cells incorporate elevated OP-P and EU levels. We speculate that these early differentiated cells exhibit a plastic phenotype and can return to a CSC state if necessary. Nevertheless, our findings imply that as differentiation progresses, the capacity of differentiated cells to revert their state becomes gradually limited due to the downregulation of the rDNA transcription machinery. In a subset of CRCs, continuous ablation of biosynthetic CSCs exhausts the progenitor cell pool and causes irreversible tumor cell differentiation, an observation that may inspire the design of effective anti-CSC therapies.

ACKNOWLEDGEMENTS

We thank all members of the Batlle lab and IRB Barcelona core facilities for histopathology, genomics, mouse mutant, and microscopy as well as the flow cytometry for extensive support. Raúl Méndez for advice. Antonio Berenguer and Camille Stephan-Otto Attolini for statistical analyses, Giorgio Stassi and the HUB (Netherlands) for sharing PDO7 and PDO-p18 respectively. IDIBAPS and Hospital del Mar Biobanks for samples. CM and AA held La Caixa predoctoral fellowships and JS Juan de la Cierva postdoctoral fellowship from the MINECO. HH is a Miguel Servet (CP14/00229) researcher funded by ISCIII, the AEI and FEDER (SAF2017-89109-P). CM is a Postdoctoral Fellow of the AECC. This work has been supported by grants from ERC (ERC AdvG 340176), WWCR (Grant 12_1209) and AGAUR (2017 SGR 698) to EB, and by FIS (PI16-01898) to AV.

AUTHOR CONTRIBUTIONS

EB supervised the study. EB, CM, JS wrote the manuscript. CM analyzed biosynthetic cells in tissues, EPHB2+ tumor cells, generated knock-in PDOs, performed experiments of cell ablation, tested the effects of chemotherapy in and performed lineage tracing experiments. JS analyzed the effects of WNT blockade, and analyzed 5.8S in tumor samples. XH-M performed xenograft experiments. CC and FS performed LGR5 ablation experiments. DS helped in screen KI organoids and AA in KRT20 quantification and clonal analysis. GT generated Lgr5 and Krt20 lineage tracing organoids. MS and AC performed IHC. GW generated NTCF4 cell lines. AM-S quantified Lgr5 and KRT20 ISH/IHQ in patient samples. ES provided logistic support. EM, CM and HH performed scRNAseq experiments. SL and LED generated AKP tumoroids.

DECLARATION OF INTERESTS

The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to eduard.batlle@irbbarcelona.org.

FIGURE TITLES AND LEGENDS

Figure 1. rDNA Transcription and Protein Synthesis across Different Models of CRC. (A) OP-P incorporation in a patient-derived xenograft (PDX). *Lu: Lumen; Str: Stroma.* Scale bars: in (A): 500 μm; in (A'): 100 μm; in (A"): 250 μm. Dashed lines outline epithelial glands. **(B)** OP-P incorporation in primary human CRC implanted orthotopically in mice. *Str: Stroma.* Scale bars: in (B): 2.5 mm; in (B'): 100 μm; in (B"): 250 μm. Dashed lines outline epithelial glands. **(C)** OP-P incorporation in spontaneous liver metastasis generated from intrasplenic (IS) injection of PDO7. *Lu: Lumen.* Scale bars: in (C): 5 mm; in (C') and (C"): 250 μm. **(D)** EU incorporation in a PDX. Dashed lines outline epithelial glands. *Lu: Lumen.* Scale bars: in (D): 500 μm; in (D'): 250 μm; in (D"): 50 μm. Images in (A), (B), (C) and (D) are tiled and stitched. **(E)** Analysis of OP-P and EU in EPHB2 tumor cell populations. **(F)** Representative flow cytometry plots of EU and OP-P into EPHB2 positive or negative tumor cells purified from indicated PDXs. **(G)** % of OP-P and EU positive cells within EPHB2 high and low tumor cells. ****$p=1.92x10^{-7}$, ****$p=1.51x10^{-9}$ in a mixed effects linear model (n=3 PDX, 2 mice for each PDX). **(H)** EPHB2-based cell purification from human CRC samples. **(I)** GSEA of EPHB2 high and low tumor cells sorted from primary CRCs. For all GOSLIM terms FDR q-val=0. **(J)** Spatial zonation model of protein synthesis/rDNA transcription in CRCs.

Figure 2. Characterization of Biosynthetic Tumor Cells in CRC

(A and **B)** Representative Ki67 and OP-P (A) or EU (B) patterns in PDXs. Dashed lines outline epithelial glands. Scale bars: in (A): 100 μm; in (B): 250 μm. **(C and D)** *In situ* hybridization of Lgr5 and immunohistochemistry (IHC) of KRT20 in serial sections of two representative human primary CRCs. Scale bars: 100 μm. **(E)** Quantification of Lgr5+, KRT20+ and double negative cells across a collection of primary CRC human samples (reproduced from Merlos-Suarez et al., 2011). **(F)** Schematic representation of the Lgr5-inducible Caspase9-tdTomato (LiCT) targeting construct. **(G** and **J)** RT-qPCR analysis of ISC and differentiation genes in PDO-p18-LiCT (G) and PDO7-LiCT (J). Bars depict the mean and upper/lower limits of relative expression obtained from representative experiment. **(H** and **K)** LGR5-tdTomato staining in tumor xenografts. Scale bars: in (H): 2.5 mm; in (K): 1 mm. **(I** and **L)** LGR5-tdTomato and OP-P staining in xenografts. Scale bars: in (I): 50 μm and in (L): 100 μm. **(M, N,** and **O)** OP-P and KRT20 staining in PDO7 xenografts. In **(M)** dashed line delimits a KRT20 positive gland in contact with the stroma. **(N)** and **(N')** show KRT20+ cells in the lumen of a tumor gland. **(O)** Single KRT20+ cells intermingled within the OP-P positive domain (arrows). *Str: stroma; Lu: Lumen*. Scale bars: in (M): 250 μm; in (N): 500 μm; in (O): 100 μm. Images in (A), (B),(I),(L),(M),(N) and (O) are tiled and stitched.

Figure 3. WNT-Driven Differentiation Causes Protein and RNA Synthesis Decrease. (A) qRT-PCR analysis of ISC and differentiation genes in control (Wnt ON) or differentiated (Wnt OFF) LS174T CRC cells**. (B)** qRT-PCR analysis of ISC and differentiation genes in control (APC OFF) or differentiated (APC ON) AKP mouse organoids. Bars depict the mean and upper/lower limits of relative expression obtained from a representative experiment. **(C)** OP-P incorporation and KRT20 in LS174T CRC cells over differentiation. Scale bars: 50 μm. **(D)** % of OP-P and EU positive cells detected by flow cytometry in LS174T after 7 days of *in vitro* differentiation. For OP-P ****p<0.0001 (n=4); for EU **p=0.0078 (n=3) in an unpaired two-tailed t-test. Confidence intervals (CI) are mean±SEM. **(E)** HPG incorporation and KRT20 staining in AKP mouse organoids over differentiation. Scale bars: 50 μm. **(F)** % of HPG and EU positive cells detected by flow cytometry in AKP organoids over differentiation. For HPG ****p=<0.0001 (n=4); for EU *p=0.0131 (n=3) in an unpaired two-tailed t-test. CI are mean±SEM. **(G)** pre-rRNA levels in LS174T and SW403 cells after 4 days of differentiation. Bars depict the mean and upper/lower limits of relative expression obtained from a representative experiment. **(H)** 5.8S rRNA and KRT20 staining in LS174T cells over differentiation. Scale bars: 20μm. **(I)**

Representative 5.8S and KRT20 staining of a primary human CRC sample. Scale bar: 1mm. **(J)** Representative plot of KRT20/5.8S intensity analysis from (I). KS-test $p=8.35x10^{-5}$. **(K)** Experimental set-up for (L). **(L)** HPG incorporation and KRT20 staining in LS174T cells during differentiation. Arrows pointing at double positive cells on day 4. Scale bars 20 µm. Images are tiled and stitched. **(M)** Experimental design for protein tracing experiment. **(N)** Immunoprecipitation of KRT20 using whole protein extracts from HPG-labeled and traced samples. Numbers below the panels indicate intensity of the signal. **(O** and **P)** Quantification of KRT20 and HPG intensity at day 3 (O) and at day 7 (P) of differentiation.

Figure 4. POLR1A High Tumor Cells Display high Biosynthetic and Tumorigenic capacities. (A) CRC cell line differentiation via inducible NTCF4. **(B)** RT-qPCR of rRNA transcriptional machinery components in LS174T after 4 days of *in vitro* differentiation. Bars depict the mean and upper/lower limits of relative expression obtained from representative experiment. **(C)** Western blot analysis of POLR1A and KRT20 in LS174T during differentiation. **(D)** POLR1A staining in LS174T tumor cells after 7 days of *in vitro* differentiation. Scale bars: 20 µm. Right panel: quantification of nuclear area occupied by POLR1A in control (Wnt ON) or upon differentiation (Wnt OFF). p=0.0005 in unpaired two-tailed t-test (n=4 images for Wnt ON and n=3 images for Wnt OFF). CI are mean±SEM. **(E)** Knock-in construct for the generation of EGFP-POLR1A fusion protein. **(F)** EGFP-POLR1A labels the nucleoli of PDO-p18 *in vitro*. Scale bar: 50 µm. **(G)** EGFP-POLR1A and KRT20 staining in PDO-p18 xenografts. Scale bars: 50 µm, and 10 µm in insets. **(H)** Representative flow cytometry plot of EGFP-POLR1A in knock-in PDO-p18 xenografts (right plot). Squares indicate the sorted EGFP-POLR1A populations. Top 10% was considered to be POLR1A high. **(I)** Western blot of POLR1A and KRT20 in indicated sorted populations from (H). **(J)** Representative flow cytometry plot of EU incorporation in EGFP-POLR1A cells purified from PDO-p18 xenografts. **(K)** % of EU positive cells in POLR1A sorted cells from tumor xenografts. ****p<0.0001 in a mixed effects linear model (n=2 different PDX and 2 experimental replicates). **(L)** Representative flow cytometry plot of OP-P incorporation in EGFP-POLR1A tumor cells populations purified from PDO-p18 xenograft**. (M)** % of OP-P positive cells in POLR1A sorted cells from tumor xenografts. ***p=0.001 in a mixed effects linear model (n=2 different PDX, 3 replicates for PDO-p18 and 1 for PDO7)**. (N)** OP-P and POLR1A (left panel) and KRT20 and POLR1A (right panel) staining in PDO-p18 xenograft. Scale bar: 50 µm. **(O)** Tumor initiation capacity of EGFP-POLR1A high and low tumor cells purified from PDO-p18 (left) and PDO7 (right). Kaplan-Meier

plots (n=8 xenografts). **p=0.0032 for PDO-p18 and ****p<0.0001 for PDO7 in a Gehan-Breslow-Wilcoxon test.

Figure 5. Mapping of Tumor Cell Populations by Single Cell profiling and clonal analysis. (A and **F)** T-SNE plots of 388 (PDO-p18, (A)) and 511 (PDO7, (F)) single-cells from tumor xenografts colored by cluster identities. **(B, C, G** and **H)** Normalized expression of the LGR5 and of the POLR1A high signatures in PDO-p18 (B and C) and PDO7 (G and H). **(D, E, I** and **J)** Trajectory graph representations for PDO-p18 and PDO7, color-coded by pseudo-time ordering of cells (D and I) and their cluster identity (E and J). **(K)** Knock-in constructs and experimental design for lineage tracing experiments. **(L)** Quantification of total number of clones per tumor area at indicated time points. Each dot represents quantifications in a section of a different tumor. n=4-10 tumors per time point and cell population.*p=0.0107, **p<0.002 in a two-way ANOVA using Turkey's multiple comparison test. CI are mean±SEM. **(M)** Quantification of number of large clones (>15 cells) per tumor area at indicated time points. Each dot represents quantifications in a section of a different tumor. n=4-10 tumors per time point and population. ****p<0.0001 in a two-way ANOVA using Turkey's multiple comparison test. CI are mean±SEM. **(N)** Quantification of clone size over time. Data represents the average number of cells per clone. Each dot represents quantifications in a section of a different tumor. n=4-10 tumors per time point and population. *p=0.026 and ***p=0.005 in a one-way ANOVA using Turkey's multiple comparison test for the last time points. CI are mean±SEM. **(O)** Quantification of the total number of tomato positive cells per tumor sample over time. Data represents the average of tomato positive cells per time point. Each dot represents quantifications in a section of a different tumor. n=4-10 tumors per time point and population. *p=0.035 and ***p=0.007 in a one-way ANOVA using Turkey's multiple comparison test for the last time points. CI are mean±SEM. **(P)** Representative images of tdTomato on tumor xenografts. Arrowheads point to single cell clones induced at day 3. Scale bars: 50 μm in d3, 100 μm in d15 and d23/30. **(Q)** IF for KRT20 and tdTomato in tumor sections after 23d or 30d of tamoxifen treatment. White arrowheads indicate tdTomato clones containing KRT20+ cells. Scale bars: 50 μm. Images in (Q) are tiled and stitched.

Figure 6. POLR1A-High Tumor Cells sustain CRC Growth.

(A) POLR1A knock-in cell ablation cassette (POLiCT) in PDOs. **(B)** Experimental design of *in vivo* ablation experiments. **(C** and **D)** Representative images of tdTomato and KRT20 in xenografts from POLiCT knock-in PDOs. Scale bars: 250 μm. **(E)**

Representative flow cytometry plots of tdTomato expression in controls or 4 days-post POLR1A ablation in PDO7 (left) and PDO-p18 (right) *in vivo.* WT unmodified clones are shown for reference. **(F)** Tumor growth during chronic POLR1A ablation. ****p<0.0001, ***p=0.004 in a two-tailed t-test. For PDO7: n=3 control and n=6 treated tumors. For PDO-p18 n=5 control and n=6 treated tumors. CI are mean±SEM. **(G)** OP-P incorporation in control tumors or after 10 days of POLR1A ablation. *Str: stroma*. Dashed lines delimit tumor glands from stroma. Scale bars: 250 µm and 100 µm. **(H)** KRT20 staining in control or POLR1A-ablated PDO7. Scale bars: 2.5 mm. **(I)** Quantification of KRT20 positive areas in control or POLR1A ablated PDO7 tumors. ***p= 0.0009 in an unpaired two-tailed t-test. (n=8 controls and n=9 treated tumors). CI are mean±SEM. **(J)** Knock-in ablation cassette for the generation of LGR5-iCaspase9-tdTomato (LiCT). **(K)** Representative flow cytometry plot of LGR5-tdTomato levels in control or after 5 days of dimerizer in PDO-p18. **(L)** Tumor growth during chronic ablation of LGR5+ tumor cells in PDO-p18. ***p=0.0010 **(M)** OP-P patterns in untreated mice or after 14 days of dimerizer. Dashed lines delimit the tumor glands. Scale bars: 250 µm. *Str: stroma.* **(N)** Knock-in ablation cassette for the generation of LGR5-iCaspase9-EGFP (LiCG) in PDO7. **(O)** Representative flow cytometry plot of LGR5-EGFP in control conditions or after 5 days of dimerzer in PDO7. **(P)** Tumor growth during chronic ablation of LGR5 tumor cells in PDO7. **(Q)** OP-P patterns in untreated mice or after 14 days of dimerizer. Scale bars: 250 µm. Images in (C) and (D) are tiled and stitched.

Figure 7. Loss of Biosynthetic Capacity drives CRC into an Irreversible Differentiation State. (A) Experimental design. **(B)** Tumor growth in untreated, during permanent dimerizer treatment (in red), or after 10 days of dimerizer treatment (in yellow) to ablate POLR1A in PDO7. ***p=0.0010 ***p=0.0008 in an unpaired two-tailed *t*-test (n=3 control, n=6 always dimerizer, n=6 stop dimerizer tumors). CI are mean±SEM. **(C)** OP-P incorporation and KRT20 expression in control tumors or in tumors 7 days after dimerizer withdrawal. Scale bars: 1mm (upper images) and 2.5 mm (lower images). **(D)** Experimental design for *in vitro* reversibility experiments in AKP organoids. Arrows indicate sample collection **(E** and **F)** Quantification of EU (E) or (F) HPG incorporation during differentiation (ON) and after reversion (ON/OFF). For EU: p=0.6, p=0.5, **p=0.004, **p=0.0019 (n=4, 3, 2, 3, 3) and HPG: p=0.62, **p=0.007, ****p<0.0001, **p=0.0015 (n=4, 4, 2, 4, 2) in an unpaired two-tailed t-test. CI are mean±SEM. **(G)** Organoid formation assay during *in vitro* differentiation (ON) and after reversion (ON/OFF). p=0.10, *p=0.01, p=0.06 (n=4, 4, 4, 2) in an unpaired two-tailed *t*-test. CI are mean±SEM. **(H)** Representative examples of POLR1A and

KRT20 staining in organoids during *in vitro* differentiation (ON) or after reversion (ON/OFF). Scale bars: 50 μm. **(I)** Quantification of POLR1A nuclear area from (H). p=0.6, *p=0.01, ***p=0.0001, ***p=0.0006 (n=6, 4, 3, 5 images per condition) in an unpaired two-tailed t-test. CI are mean±SEM. **(J)** Experimental design for *in vivo* reversibility experiments of AKP tumors. **(K)** Tumor growth corresponding to (J). **p=0.0088, p=0.709 in an unpaired two-tailed t-test (n=8, 10, 7 tumors). CI are mean±SEM. **(L)** Representative images of HPG incorporation and KRT20 in AKP tumor xenografts. **(M)** POLR1A and KRT20 in AKP xenografts during differentiation (ON) and after the reversion (ON/OFF). Scale bars: 50 μm **(N)** Differentiation-driven irreversible loss of plasticity in CRC. Images in (H), (L) and (M) are tiled and stitched.

STAR METHODS

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Eduard Batlle (eduard.batlle@irbbarcelona.org).

Materials availability

Knock-in CRC organoids and NTCF4 expressing CRC cell lines generated in this study are available from the Lead Contact with a completed Materials Transfer Agreement.

Data and code availability

Expression arrays and RNA-seq data are available at GEO with accession number: GSE125232. scRNA-seq data is available at GEO with accession number: GSE148345. Computer code is available upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Mouse strains and tumor cell injections

All experiments with mouse models were approved by the Animal Care and Use Committee of Barcelona Science Park (CEEA-PCB) and the Catalan government. To generate subcutaneous xenografts from PDOs we inoculated 150.000 (PDO7), 50.000 (PDO p19b) or 1.5 million cells (PDO-p18) in a format of 4-5 days grown

undissociated organoids into Beige (CB17.*Cg-PrkdcscidLystbg-J/Crl*) or NSG (NOD.*Cg-Prkdc<scid>Il2rg<tm1Wjl>/SzJ*) female mice in 100 µl of 50% BME2-HBSS (Cultrex). For mouse-derived AKP xenografts, 5000 cells in 100µl of 50% matrigel-HBSS (Corning) were injected into the flanks of Swiss nude (*Crl:NU(Ico)-Fox1nu*) female mice. Cell number was determined by trypsinizing and counting an aliquot of a total cellular suspension. Tumor volume was measured with manual calipers and calculated according to the formula (length x width x height)/2. To generate liver metastasis dissociated organoids (single cells) were injected intrasplenically (IS) in NSG mice as described in (Calon et al., 2012; Tauriello et al., 2018). Cells were counted and resuspended in HBSS for injection using 0.5 x 10^6 cells in 70uL per mouse. For tumor initiation assays, viable single cells from disaggregated xenografts were sorted according to their EGFP-POLR1A levels and transplanted subcutaneously into Beige/NSG recipient mice in 100uL of 50% BME2-HBSS. For metastasis initiation experiments, viable single cells from disaggregated xenografts were FACS sorted according to their POLR1A-tdTomato levels and transplanted intrasplenically in NSG mice in 70ul of HBSS.

***In vivo* treatments**

For the inducible ablation experiments, we treated animals with 5mg/kg of dimerizer (AP20187, Medchem express, HY-13992) when tumors reached an average size of 40-50 mm^3. For acute treatments (5-10 days) animals were treated daily and for chronic treatments (>10 days) every 2-3 days. Control animals were treated with vehicle for the same period of time. For *in vivo* reversibility experiments, mice bearing subcutaneous AKP xenografts were given 2 mg/ml DOX in drinking water until the tumors reached the size of about 50mm^3. After that, DOX was removed for the indicated periods (8 days ON or 20 days APC-ON) before the mice were given DOX again (APC-OFF). To induce Cre-mediated recombination of the tdTomato cassette for the linage tracing experiments, mice bearing subcutaneous tumor xenografts (50-70mm^3) receive a single intraperitoneal tamoxifen (Sigma) injection of 1mg/kg for KRT20 and 10mg/kg for POLR1A or two consecutive intraperitoneal tamoxifen injections of 250mg/kg for LGR5. For the experiments with the small molecule RNA polymerase I inhibitor, BMH21 (Invivochem), we injected intraperitoneally a daily dose of 100 mg/Kg over the period of 10 days once the tumors have reached an average size of 50 mm^3. Control animals were treated with vehicle over the same period. For chemotherapy experiments subcutaneous tumor xenorgrafts were generated by inoculating POLI-iCT PDO7 or PDO-p18 knock-in organoids. Once the tumors reach an average size of 70mm^3 animals were treated

with three intraperitoneal injections of folfiri (100ul/20gr) at day 0, 3 and 7. Folfiri cocktail contained 5-FU (30mg/kg), Irinotecan (24mg/kg)) and Leucovorin (90mg/kg). Control animals were treated with vehicle following the same regimen. Animals were sacrificed the last day of the treatment and tumors were removed and processed for FACS and histological analysis. In all experiments animals were sacrificed at indicated time points or just before tumor size reached the volume permitted by Animal Facility regulations. In all animal treatments, individuals were randomized.

Orthotopic models and primary CRC samples

Patient-derived orthotopic xenografts (PDOX1 and PDOX2) were generated from fresh biopsies taken of peritoneal implants at tumor relapse after FOLFOX chemotherapy treatment from Hospital Universitari de Bellvitge (L'Hospitalet de Llobregat) and Hospital Moisès Broggi (Sant Joan Despi) from Barcelona. Orthotopic implantation procedure of colorectal human tumors was performed as described previously (Aytes et al., 2012). At the time of sacrifice mice implanted with PDOX2 tumor presented a synchronic hepatic metastasis that was isolated and implanted orthotopically into the liver of other four animals (PDOX2-MH). All patients gave written consent to participate in the study, and The Ethics Committee of the hospitals approved the study protocol. Animals were kept under pathogen-free conditions, and all animal work was conducted according to the guidelines from the Animal Care Committee at the Generalitat de Catalunya (Procedure 9111) in the Animal Core Facility of IDIBELL (AAALAC Unit 1155). The study was approved by the IDIBELL Ethical Committee and performed in accordance with guidelines stated in The International Guiding Principles for Biomedical Research involving Animals, developed by the Council for International Organizations of Medical Sciences (CIOMS). For the purification of EPHB2 tumor cell populations, fresh surgically resected Human CRC samples (n=10) were obtained from individuals treated at Hospital Clínic (Barcelona). Tumors were disaggregated following the protocol detailed in (Calon et al., 2012) and EPHB2 populations were purified as described in (Merlos-Suárez et al., 2011). For the 5.8S and KRT20 analysis, formalin-fixed, paraffin-embedded tissue blocks of primary human colorectal adenocarcinomas (n=24) were retrieved from the biobank of the Servei de Patologia from Hospital del Mar (Barcelona). Samples were obtained under informed consent and approval of the Tumor Bank Committees according to Spanish ethical regulations. The study followed the guidelines of the Declaration of Helsinki, and patient identity for pathological specimens remained anonymous in the context of this study.

Organoid *in vitro* culture

Patient-derived organoids (PDOs) were cultured as described in (Cortina et al., 2017; van de Wetering et al., 2015). PDO7, was obtained from Prof. Stassi laboratory (Lombardo et al., 2011), PDO-p18 and PDO-p19b from Prof. Clevers laboratory (van de Wetering et al., 2015) and AKP (shAPC-KP) organoids were a kind gift from Dr. Luke Dow's laboratory. AKP organoids were embedded in Matrigel (Corning) and grown in Advanced DMEM/F12 supplemented with 10 mM HEPES, 2 mM GlutaMAX, B27, recombinant Noggin (100 ng/ml), EGF (50 ng/ml, Preprotech), and in the presence of Doxycycline (DOX, 2 µg/mL), (APC OFF) or absence (APC ON) to induce differentiation (Dow et al., 2015) at indicated time points. PDOs were embedded in a mixture of 70% BME2 and 30% Advanced DMEM/F12 with HEPES and GlutaMAX. PDO-p18 and PDO-p19b were cultured as is described by the Clevers laboratory (van de Wetering et al., 2015) and PDO7 as described in (Cortina et al., 2017).

N-TCF4 expressing CRC cells lines

LS174T and SW403 CRC cell lines obtained from ATCC were engineered to express a tamoxifen-inducible dominant negative form of TCF4 which consists in the β-catenin-binding domain of TCF4 fused to a modified hormone-binding domain of the estrogen receptor (nTCF-ERT2)(van de Wetering et al., 2002). CRC cell lines expressing the nTCF-ERT2 construct were cultured under standard conditions (DMEM with 10% FBS) (Wnt ON) and when indicated, were treated with 1 µM 4-hydroxytamoxifen (Merck H7904) (Wnt OFF). In all experiments, we only used the modified cell lines, and for simplicity we refer to them as LS174T and SW403 respectively, although both express nTCF-ERT2 construct.

METHOD DETAILS

Organoid formation experiments

To run organoid formation experiments we used single cells obtained either by trypsinization from an *in vitro* experiment, or by FACS sorting from subcutaneous tumor xenografts. For the AKP organoid formation experiments, AKP organoids embedded in Matrigel were grown in the presence of DOX (APC OFF) or without DOX (APC ON) for the indicated periods (2 days ON, 5 days ON). Then, organoids were disaggregated to a single cell suspension using TrypLE reagent (Gibco), and seeded in 25 μL Matrigel drops containing 1000 cells/drop, and DOX was added to

the medium in all conditions indicated with /OFF. To assess POLR1A organoid formation capacity, we disaggregated xenografts as described below and living cells were selected for EPCAM expression and sorted according to their endogenous POLR1A levels by FACS. Afterward, single cells were seeded *in vitro* in 25 µl BME2 drops containing 1000 cells/drop per condition. To test the effect of BMH21 *in vitro*, single cells obtained by trypsinization from *in vitro* culture were seeded in 25 µl BME2 drops containing 1000 cells/drop per condition and were treated daily with BMH-21 (1µM). Control plates were treated with a vehicle and the growth was measured using the below described method. In all cases, plates were scanned using ScanR inverted microscope (Olympus) at day 1 post-seeding to quantify the exact number of cells seeded per drop and at the experimental endpoint (day 7 post-seeding). Full drops were scanned taking overlapping pictures at 4x magnification and at 8 different z-stacks with a separation of 200 µm. Z-stacks of all images were projected into a single image and the full drop was digitally reconstructed by stitching the different image projections using an Image J (Schindelin et al., 2009) custom-made macro developed for this purpose. Total number and mean size of cells (i.e. any object with a diameter larger than 5 µm) or organoids (diameter larger than 400 µm) were counted. Differences were assessed with Student's *t*-test.

Immunostaining and confocal imaging

CRC cell lines grown on ethanol-sterilized glass coverslips were washed three times with PBS and incubated with 4% paraformaldehyde (PFA) for 15 min at room temperature, washed again with PBS and incubated for additional 5 min with 50 mM NH_4Cl to quench PFA's auto-fluorescence. Blocking and permeabilization were performed simultaneously for 1 h at room temperature in PBS containing 3% BSA and 0.3% Triton X-100. Coverslips were first incubated in a humid chamber overnight at 4ºC with the primary antibody diluted in the blocking/permeabilization solution. After extensive washing, the incubation with the secondary antibody was done for 1 h in the dark. In the case of the co-staining, both primary and secondary antibodies were mixed and used at the same time. Coverslips were mounted with DAPI-Fluoromount-G (SouthernBiotech. 0100-20). For live imaging, PDOs were seeded in 10 µl drops of BME2 in 8-well chamber coverslip (IBIDI) and covered with 200uL of the corresponding culture medium. 4-5 days grown organoids were directly incubated 30 min at 37ºC with Hoechst 33342 1:1000 (Molecular Probes, R37605) prior to confocal visualization. For immunostainings, PDOs were seeded in microscopy chamber slides in thin layers of BME2 (20 µL per chamber). Samples were then fixed with 4% PFA for 15 min and quenched with 20 mM glycine for 20 min.

Permeabilization was achieved with 0.5% Triton X-100 for 30 min at room temperature, then a second blocking step with 1% BSA for 1 h at room temperature was performed. Samples were incubated with primary antibodies overnight at 4ºC in a dark chamber followed by the incubation with the correspondent secondary antibody for 1 h at room temperature in the dark. Washes were performed in between each step with PBS. Coverslips were mounted with DAPI-Fluoromount-G. To visualize OP-P (Medchem Source LLP, JA-1024), EU (Click Chemistry Tools, 1261-100), and HPG (Click Chemistry Tools, 1067-100), slides were incubated with Click-it cocktail kit (Thermo Fisher) for 30 min in dark prior to antibody staining. Primary and secondary antibodies used for *in vitro* immunostaining were the same as the ones described for IF in paraffin sections. For paraffin-included organoid pellets, Matrigel drops were washed twice with PBS, incubated with formalin for 2h at RT, washed again with PBS, and left ON in formalin. The next day, two more washes with PBS were performed, and drops were put into a microcassette (Leica). Microcassettes were put into a regular histological cassette before paraffin embedding. Immunostainings were performed as is described in histology and tissue staining section. Images were taken with a Zeiss LSM780 confocal microscope. Whole organoid z-stacks were taken with a LEICA SP5 microscope. Next, images were background-substracted and maximum intensity projections were rendered by Imaris software version 9.1.

Histology and tissue stainings

Immunostainings were performed on 4 µm sections following a standard protocol. For PAS/Alcian Blue staining, samples were deparaffinated and rehydrated, followed by 5 min incubation with Alcian Blue solution pH 2.5 (1 g of Alcian Blue (Panreac), 97 ml miliQ water, 100 ml glacial acetic acid) and dehydrated with ethanol 96%. Samples were rehydrated again, washed with distilled water, incubated with 1% periodic acid for 20 min, washed with tap water, incubated for 20 min with Schiff's reagent (Sigma), washed with distilled water, and counterstained with hematoxylin prior to dehydrating and mounting with DPX (Panreac, 255254.1608). For immunohistochemistry (IHC), antigen retrieval was carried out by boiling in Tris-EDTA buffer for 20 min, then samples were blocked with Peroxidase-Blocking Solution (Dako: S202386) followed by blocking with normal goat serum 10% (Jackson Immunoresearch, 005-000-121) in wash buffer 1X and incubated with primary antibody against TdTomato (rabbit, Rockland, 600-401-379) 1:200 in Envision FLEX antibody diluent (Agilent, K8006) at 4ºC overnight. Secondary antibody BrightVision poly-HRP anti- Rabbit (Immunologic, DPVR-110HRP) was

added for 30 min at room temperature. Slides were counterstained with hematoxylin prior to dehydrating and mounting with DPX (Panreac, 255254.1608). Washes were performed in between steps with wash buffer 1X (Dako, K800721). To visualize OP-P, EU, and HPG, slides were incubated with Click-it cocktail kit (Thermo Fisher, 100 µl per slide) for 30 min in dark prior antibody staining. For immunofluorescence the applied protocol was the same as for IHC but blocking with normal donkey serum 10% (Jackson Immunoresearch, 017-000-121) in wash buffer 1X with the following primary antibodies diluted in Envision FLEX antibody diluent (Agilent, K8006): 1:50 EPHB2 (goat, R&D systems, AF467), 1:50 EPCAM-Trop1 (goat, R&D, AF960), 1:100 KI67 (mouse, BD, 550609), 1:50 5.8S (mouse, Thermo Fisher, MA1-13017), 1:100 POLR1A (mouse, Santa Cruz, sc-48385), 1:200 tdTomato or RFP (goat, Sicgen, AB8181-200), 1:200 EGFP (rabbit, abcam, ab183734), 1:50 KRT20 (mouse, Dako, M7019), and secondary antibodies: donkey anti-goat conjugated to Alexa-488/568/647 (Life Technologies A11055, A11057, A21447), donkey anti-rabbit conjugated to Alexa-488/568/647 (Life Technologies A21206, A10042, A31573) and donkey anti-mouse conjugated to Alexa-488/568/647 (Life Technologies A-21202, A10037, A31571) at 1:400 for 1h art RT. DAPI was added at 1:2500 after secondary antibody incubation and slides were mounted with Fluorescent mounting medium (Dako, 53023). For in situ hybridization paraffin sections of human tissues were de-waxed and hydrated following standard procedures. Samples were then treated with 0.2N HCl for 15min at room temperature, washed 3X in PBS and incubated for 20min at 37ºC in Proteinase K (30 µg/ml in PBS). 0,2% glycine in PBS was added for 3 minutes to neutralize Proteinase K activity and samples were then washed 2X in PBS. Sections were postfixed in 4% PFA for 10 min and washed 3X in PBS. Histone acetylation was performed by incubating the samples 2X 5 min in an H_2O solution containing 1.5% triethanolamine, 0.15% HCl and 0.6% Acetic anhydride. Samples were then washed and prehybridized for 1h at 70ºC in prehybridization solution (50% Formamide, 5X SSC, 2% Blocking Reagent (Roche), 0.05% CHAPS, 5mM EDTA, 50µg/ml Heparin and 50µg/ml yeast RNA). For Lgr5 ISH, mRNA probe was diluted 500ng/ml in prehybridization solution and incubated for 48h at 65ºC.. Post-hybridization washes were performed 3X 20 min in 50% Formamide / 2XSSC at 65ºC. Sections were then rinsed in TBST buffer (0.1M Tris-HCl pH7.5, 0.15M NaCl, 0.1% Tween20) and blocked for 30 min at room temperature in Blocking buffer (0.5% Blocking Reagent, 10% sheep serum in TBST). Sheep anti-DIG antibody (Roche) was diluted 1/5000 in blocking buffer and incubated overnight at 4ºC. Finally, samples were washed in TBST and then in NTM buffer (0.1M Tris pH9.5, 0.1M NaCl, 0.05 M $MgCl_2$) and developed in NBT/BCIP solution (Roche) for 24-72h.

Histological Image Acquisition

Digital scanned brightfield and fluorescent images were acquired with a NanoZoomer-2.0 HT C9600 scanner (Hamamatsu, Photonics, France) equipped with a 20X objective and using NDP.scan2.5 software U10074-03 (Hamamatsu, Photonics, France). Fluorescent images were acquired with a mercury lamp unit L11600-05 couple to the NanoZoomer. All images were visualized with the NDP.view 2 U123888-01 software (Hamamatsu, Photonics, France) with a gamma correction set at 1.8 in the image control panel. In each batch of samples, the same exposure time and gain per antibody have been used. All images were visualized with a gamma correction set at 1.0 and the sharpen filter enabled in the image controls panel of the NDP.view 2 U12388-01 software.

sgRNA design

Small guide RNAs were designed <15bp around de intended site of knock-in insertion using the http:// crispr.mit.edu web tool and were cloned into a pX330-IRFP hSpCas9 plasmid (Cortina et al., 2017) as described in *http://www.genome-engineering.org/crispr/wp-content/uploads/2014/05/CRISPR-Reagent-Description-Rev20140509.pdf*. The sgRNA sequences are available in Table S1.

Donor plasmid construction

750bp (LGR5, KRT20, POLR1A and AAVS1 constructs) of 5' and 3' homology arm (HA) flanking the knock-in insertion cassettes were amplified from PDO7 gDNA or synthetized by gene synthesis (Thermo Fisher) and cloned into pShuttle or pDONOR vectors. Homology arms EGFP-linker, LF2A-CreERT2-BGHpA, Lox-tagBFP-3xpA-Lox-tdTomato-BGHpA insertion cassettes were generated by gene synthesis (Thermo Fisher) and cloned in the 5'HA-3'HA previously engineered pShuttle or pDONOR vectors (as described in Cortina et al., 2017).

For the *LGR5-iCaspase9-T2A-tdTomato* donor (pDONOR-LGR5-iCT), a previous described p-shuttle-LGR5-LF2A-CreERT2 plasmid (Cortina et al., 2017) was modified by replacing the existing LGR5 5' homology arm with a LGR5 5' homology arm containing a STOP codon using PacI and SalI. Subsequently, the plasmid was digested with SgrDI/ and NotI-HF to substitute the LF2A-CreERT2 cassette with an IRES-iCasp9-T2A-TdTom-WPRE cassette, amplified by PCR from a plasmid obtained from T.Sato's laboratory (Shimokawa et al., 2017). For the *LGR5-iCaspase9-T2A-EGFP* version TdTomato was then replaced by EGFP by digesting

with KflI and MluI, and introducing in-frame an iCasp9 fragment-T2A-EGFP, amplified by PCR from a custom gene synthesis DNA string (Thermo Fisher).
For *POLR1A-iCaspase9-T2A-tdTomato,* the human POLR1A 5' homology arm was cloned in the place of the LGR5 5' homology arm in pDONOR-LGR5-iCT donor vector by PacI and SgrDI digestion. The 3' homology arm was cloned by NotI-HF and AscI digestion. Inserts were amplified by PCR from custom gene synthesis DNA strings (Thermo Fisher).

Generation of CRISPR-Cas9 knock-in PDOs

PDOs were cultured and CRISPR-Cas9 knock-in as described in (Cortina et al., 2017). In brief, to obtain PDOs expressing endogenous EGFP-POLR1A fusion protein we designed a sgRNA targeting a sequence flanking the TSS (or transcription start site) of the *POLR1A* locus. For the lineage tracing and the inducible caspase9 knock-in generation, sgRNA sequences were designed to bind the 3'UTR. After nucleofection with guide-Cas9 and donor plasmids (Cortina et al., 2017), single cell clones were derived and PCR screened for correct integration. Southern blotting was performed to assess off-target integrations. For each knock-in construct, we evaluated that the fluorescent cassette reported the expression of the gene of interest by RT-qPCR *in vitro* and *in vivo* in a similar manner as performed in Cortina et al., 2017.

Genotyping PCR and Southern blot

Single-cell derived clones were lysed in of 10 mM Tris, 1 mM EDTA, 1 % Tween 20 and 0.4 mg/ml proteinase K lysis buffer for 1h at 55°C, and the enzyme was inactivated for 10' at 95ºC. 2 µL of the lysate was directly used in the specific integration PCR. For the 5' specific integration PCR, a forward primer binding gDNA upstream of the 5' homology arm and a reverse primer at the beginning of the inserted cassette were used. Similarly, for the 3' specific integration PCR a forward primer at the end of the inserted cassette and a reverse primer downstream of the 3' homology arm was used. Primer sequences used are shown in Table S2. For southern blot, genomic DNA was extracted using the GenElute Mammalian Genomic DNA Miniprep Kit (Sigma G1N70-1KT). 10 µg of genomic DNA was digested overnight with the appropriate restriction enzyme (Table S3) and separated on a 0.8 % agarose gel. DNA fragments were transferred by capillarity to a Hybond-N+ membrane (GE Healthcare RPN303B) overnight. Probes were generated by PCR and radioactively labeled with α-[32P] dCTP using the MegaPrime labeling kit (GE Healthcare RPN1604). Hybridization with the probe was carried out overnight at

60°C. Probes were detected using a Phosphoimager plate and the probe sequences are annotated in Table S3.

RNA extraction and quantitative RT–qPCR

For *in vitro* experiments, total RNA was extracted using TRIzol (Invitrogen) followed by RNA column purification using the RNA PureLink Kit (Ambion). Briefly, cells were scraped from cell culture dishes and homogenized by pipetting in TRIzol solution. After phase separation with chloroform, the RNA from the aqueous phase was purified with the RNA PureLink Kit and quantified by Nanodrop spectrophotometer. cDNA was produced with the High-Capacity cDNA Reverse transcription Kit (Applied Biosystems) following the manufacturer's instructions.

To analyze gene expression changes RT-qPCR was performed using 5 ng of cDNA per each real-time qPCR well. Real-time qPCRs were performed with TaqMan assays (Applied Biosystems) and TaqMan Universal PCR Master Mix in triplicates following manufacturer's instructions (Applied Biosystems 4369016). Gene expression levels were normalized using the endogenous control PPIA and B2M for each sample and differences in target gene expression were determined using SDS 2.4 or StepOne 2.2 plus software. Error bars represent standard deviation of samples performed in triplicate. The following TaqMan assays were used: EGFP (Mr04097229_mr), Tomato-BGHPA (custom made probe; F: GGGCATGGCACCGGCAGCACC, R: CCTACTTGTACAGCTCGTCCATGCC), LGR5 (Hs_00173664_m1), ASCL2 (Hs_00270888_S1), SMOC2 (Hs0159663_m1), MUC2 (Hs_03005094_m1), KRT20 (Hs_00300643_m1), POLR1A (Hs00209909_m1), POLR1B (Hs00219263_m1), POLR1C (Hs01561269_g1), RRN3 (TIF1A) (Hs_04398176_m1), pre-5'ETS-rRNA (custom), Ascl2 (Mm_01268891_g1), Lgr5 (Mm_0043889_m1), Muc2 (Mm_01276696_m1), Smoc2 (Mm_00491553_m1), and Krt20 (Mm_00508106_m1). All real-time qPCRs were normalized to B2M (Hs_99999907_m1), b2m (Mm_00437762_m1), PPIA (Hs_99999904_m1), and Ppia (Mm02342430_g1).

Protein extraction and western blot analysis

Cells were washed twice with cold PBS and scraped with lysis buffer 1:1:1 (1 mM EDTA, 1 mM EGTA, 1% SDS) supplemented with protease and phosphatase inhibitors (Sigma Aldrich) and heated at 99ºC for 10 min. Cell lysates were pipetted several times to break up gDNA and later centrifuged at 13200 rpm at 4ºC for 15 min. The supernatant was kept as the protein extract. Protein content was quantified with the Protein Assay (BioRad) based on the Bradford method. Equal amounts of protein

per sample were separated by standard SDS-PAGE and transferred to PVDF membranes (Millipore). The membranes were incubated in TBS-T (20mM Tris, 150mMNaCl, 0.2% Tween 20) supplemented with 5% milk for 30 min at RT to block unspecific antibody binding. Primary antibodies RPA194 (POLR1A) (Santa Cruz, SC-48385) 1:500, RRN3 (Sigma HPA049837) 1:500, KRT20 (DAKO M701929-2) 1:500, Actin (Abcam ab20272), 1:15000 were incubated overnight at 4ºC. Secondary antibodies were diluted 1:10000 and incubated for 1h at RT with the PVDF membranes. Membranes were washed 3 times with TBS-T 0.2% for 10 min, incubated for 5 min with SuperSignal West Pico Chemiluminescent Substrate (Thermo Scientific, 34080), and visualized using Odyssey Fc Imaging system (Li-COR).

***In vitro* click chemistry, in-gel fluorescence, IP**

Cells were lysed with 2% SDS, extracts were quantified and diluted to 1% SDS, and 100 µg of protein was subjected to click reaction using Click-it Kit (Thermo Fisher) and IRDye 800CW Azide Infrared Dye (Li-COR). Samples were incubated in dark at room temperature for 1.5 h, and subsequently, the proteins were precipitated using standard methanol/chloroform protocol. Pellets were air dried and re-suspended in 1% SDS. 20 µg of protein was boiled with sample buffer and loaded onto SDS-PAGE gel for further analysis of the inputs or was incubated with antibody against KRT20 (Atlas antibodies, HPA024309) overnight at 4°C with shaking. The next day, protein A beads were added, and samples were incubated 1 h more under the same conditions. Subsequently, after extensive washes, samples were boiled with sample buffer and loaded onto SDS-PAGE gel followed by standard WB protocol. In order to detect KRT20, membranes were incubated for 1 h in dark at room temperature with 1:5000 Goat anti-Rabbit Alexa Fluor 680 (Thermo Fisher). Membranes were washed 3 times with TBS-T 0.2% for 10 min, incubated for 5 min with SuperSignal West Pico Chemiluminescent Substrate (Thermo Scientific, 34080), and visualized using Odyssey Fc Imaging system (Li-COR).

Analysis of protein and RNA synthesis by flow cytometry and microscopy

To measure protein synthesis in vivo mice were injected intraperitoneally with either 1 mg of OP-P (O-propargyl-puromycin) (Medchem Source LLP, JA-1024) or 2 mg of HPG (L Homopropargylglycine) (Click Chemistry Tools, 1067-100) per animal two hours before sacrifice. To measure RNA synthesis in vivo mice were injected i.p. with 4 mg EU (5-Ethynyl Uridine) (Click Chemistry Tools, 1261-100) per animal two hours before sacrifice. Subcutaneous xenografts were collected and further processed for

histology or flow cytometry analysis following the standard protocols. FACS Aria Fusion flow cytometer (Beckton Dickinson) was used and data were analyzed using FlowJo software. Using a Click-It kit (Thermo Fisher, C10457, C10428, C10329) OP-P, HPG or EU were conjugated to a fluorochrome in tissue sections prior to antibody staining or in FACS sorted cells which were fixed in PFA 4% for 15 min at RT before Click-It reaction. Samples were then analyzed by microscopy or flow cytometry in each case. Samples from non-injected mice were used to determine the background signal. In order to measure protein or RNA synthesis in vitro, either OP-P (20 μM), HPG (50 μM) or EU (1 mM) was added to the medium. Cells were incubated for 2h, then trypsinized to single cell state for flow cytometry analysis, fixed, and labeled by standard Click-It reaction protocol. For microscopy analysis, cells or organoids were washed two times with PBS after the incubation period, fixed and subjected to Click-It reaction under the same conditions as stated before. When indicated, samples were stained with antibodies and further processed following standard immunofluorescence protocol.

Xenografts disaggregation and cell isolation

Xenografts were resected and disaggregated as previously described in (Merlos-Suárez et al., 2011)·(Cortina et al., 2017). Human epithelial cells from disaggregated PDX were first incubated 30 min at 4ºC with 1:200 CD16/CD32 (mouse, Tonbo Biosciences, 70-0161-U500) to block free antibody binding sites and with 1:200 BV421-CD31 (rat, BD Biosciences, 562939cloneMEC13.3) and 1:200 BV421-CD45RB (rat, BD Biosciences, 562849clone16A) to stain for immune and endothelial mouse cells. After this period, 1:150 EPCAM-PeCy7 (human, eBioScience 25- 9326-42) or 1:100 EPCAM-APC-Vio770 (human, Miltenyl Biotec 130-101-161) was added and incubated for 1 h at 4ºC. Mouse tumor cells from AKP xenografts were stained with 1:300 EPCAM-APC-Cy7 (mouse, Biolegends, 118217 cloneG8.8). DAPI (1 μg/ml) was added to distinguish live/dead cells. The cell suspension was analyzed with a BD Fusion FACS or Aria FACS.

Generation of POLR1A and LGR5 signatures

The POLR1A and LGR5 gene signatures used for the single cell RNA-seq analysis were generated by FACS purification of high and low or negative fluorescent tumor cells from subcutaneously injected POL-iCT or LGR5-iCT knock-in PDO7. The LGR5 signature included bonafide ISC genes that were corregulated with LGR5 according to the scRNAseq data. All gene signatures used in this study are provided in Table S5.

Microarrays

RNA from 2000 cells from each condition was extracted and retrotranscribed to cDNA as described in (Gonzalez-Roca et al., 2010). The cDNA was amplified for and purified using PureLink Quick PCR Purification Kit (Invitrogen). Quantification of amplified cDNA was done on a Nanodrop ND-1000 spectrophotometer (Thermo-Fisher Scientific, Waltham, MA, USA). 8 ug of the cDNA from each sample were fragmented and labeling with GeneChip Mapping 250 K Nsp assay kit (Affymetrix) following the instructions of the manufacturer. Finally, hybridization was performed using the GeneChip Hybridization, Wash and Stain Kit (Affymetrix). Samples ready to hybridize were denatured at 99°C for 2 min prior to incubation into the GeneChip Human PrimeView arrays (Affymetrix). Hybridization was performed for 16 h at 45°C / 60 rpm in the GeneChip Hybridization Oven 645 (Affymetrix). Washing and Stain steps after hybridization were performed in the GeneChip Fluidics Station 450 (Affymetrix), following the specific script for PrimeView arrays. Finally, the arrays were scanned with GeneChip Scanner GCS3000 (Affymetrix) using default parameters, and the generation of Cel files for bioinformatics analysis was done with Command Console software (Affymetrix).

Cell sorting and libraries preparation for Single-cell RNA-seq

PDO7 was infected with a CMV-CherryLuc lentiviral vector. Positively infected cells were sorted, expanded for two passages, and injected subcutaneously in a Beige nude mouse (200.000 cells in small organoid format per flank in 100µl BME2-HBSS). When PDX were at a mean $100m^3$ volume the animal was sacrificed and xenografts were disaggregated as described before. Cherry positive alive cells from two xenograft biological replicates were sorted in single cell format in 96-well plate with Smart-seq2 lysis buffer and frozen prior to their sequencing. For PDO-p18 2 million cells were injected subcutaneously. $100m^3$ xenografts were extracted from the NSG injected animal after its sacrifice, disaggregated and stained for 1h on ice with hEpcam-APC-770 antibody (human, Milteny Biotec, 130-101-161) at 1:150 dilution after 30 min of mouse antigen blocking stain with 1:100 mCD16/CD32 (Tonbo Biociences, 70-0161-U500), 1:100 mCD31-vio450 (BD biosciences, 562939cloneMEC13.3) and 1:100 mCD45-vio450 (BD biosciences, 562849clone16A). Epcam positive human cells from two different replica xenografts were sorted in single cell format 96MW plates with Smart-seq lysis buffer.

Full-length single-cell RNA sequencing libraries were prepared using the Smart-seq2 protocol (Picelli et al.) with minor modifications. Briefly, 96-well plates containing sorted cells in lysis buffer were reverse transcribed using SuperScrpit II (Invitrogen) in the presence of oligo-dT30VN, template-switching oligonucleotides and betaine. The cDNA was amplified using the KAPA HiFi Hotstart ReadyMix (Kappa Biosystems), ISPCR primer and 22 cycles of amplification. Following purification with Agencourt Ampure XP beads (Beckmann Coulter), product size distribution and quantity were assessed on a Bioanalyzer using a High Sensitivity DNA Kit (Agilent Technologies). 140 pg of the amplified cDNA was fragmented using Nextera® XT (Illumina) and amplified with indexed Nextera® PCR primers. Products were purified twice with Agencourt Ampure XP beads and quantified again using a Bioanalyzer High Sensitivity DNA Kit. Sequencing of Nextera® libraries from 384 cells was carried out in one lane on an Illumina HiSeq2500 v4.

QUANTIFICATION and STATISTICAL ANALYSES

General quantifications

For the ribosomal content/KRT20 ratio quantification in primary patient samples, scanned images were processed for quantification using Qupath software (https://qupath.github.io/). The epithelial area was selected and fragmented into 100 μm^2 tiles. Mean intensity for the red and green channel was measured for each tile, and the values were represented using a scatter plot. Consecutive slides for 8 representative samples were stained for EPHB2 and 5.8S as a control. Densities of absolute-valued Log2 ratios were used to perform statistical analysis.

For quantifying the number of KRT20 and tdTomato positive cells for the KRT20 lineage tracing experiment, images were processed using Qupath as described above and the Cell detection algorithm from the same software was used. Nuclei were detected by using the DAPI channel with default nuclei parameters and tdTomato in the 568 channel. The mean intensity of KRT20 staining was measured for each tdTomato positive detected cell.

For Ki67 quantification images were processed for quantification using Qupath. In order to compare the amount of KI67 positive cells in tumor center compared to tumor periphery, we selected 4 areas within each image: a center, and three zones representing inner (P3), outer (P2), and the outermost periphery (P1). Epithelial tumor cells were classified as positive or negative based on mean DAB nuclear intensity, and percentage of positive tumor cells within each zone was represented.

ImageJ was used for the quantification of the whole silde images of tissue sections stained for KRT20 or alcian blue. For KRT20 quantification, epithelial tumor cells were first selected as EPCAM+ and then we quantified the KRT20+ area within the epithelial compartment. Percentage of Alcian blue area represents a ratio between Alcian blue+ area and the total area of the tumor.
The quantification of the nuclear area occupied by POLR1A signal was done with ImageJ using a Macro developed by the microscopy facility at IRB.

Clonal quantification for *in vivo* lineage tracing

We tested several tamoxifen doses and assessed the % of recombination by flow cytometry in dissociated xenograft. A tamoxifen dose of 250mg/Kg in LGR5-creERT2 PDOs, 10mg/Kg in PORLR1A-creERT2 and 1mg/Kg in KRT20-creERT2 trigger recombination in 1-2% of the tdTomato+ cells in the viable epithelial fraction. We corroborated this result by histology. Mice were inoculated subcutaneously with 200.000 PDO cells and treated with the indicated doses of tamoxifen when xenografts reached 50-70 mm3 (approximately 3 weeks after inoculation). Tumors were processed for histology at indicated time points. A minimum of 4 tumor xenografts per cre line were analyzed. Tumors were open in half and oriented so that sections corresponded to the central area of the tumor. When all samples were collected, we performed immunohistochemical staining of 3 µm histological sections as described above using the primary antibody against tdTomato (rabbit, Rockland, 600-401-379). Images for clone quantification were acquired using a NanoZoomer-2.0 as described above. All brightfield images were visualized with a gamma correction set at 1.8 with the NDP.view 2U123888-01 software (Hamamatsu). For clonal quantification images were processed using ImageJ. Tomato positive objects were detected and thresholds were calculated automatically using RenyiEntropy function. The algorithm group objects that reside in close proximity (<13 pixels) so that they are considered as unique clones. Total and large clones number were relativized to the area (pixels) of the tumor section. To calculate number of cells, we estimated the size of single cells in histological sections and interpolated the number of cells per each clone and per section. The non-necrotic area include all tumor cells (epithelium and stroma) that were not necrotic and was obtained by manually training a machine learning tile classifier using QuPath software. This area was then divided by the total tumor area in order to obtain the non-necrotic fraction.

POLR1A expression analysis in colon samples from human cohort

GSE4407 dataset (Sanz-Pamplona et al., 2014) was downloaded from the NCBI GEO repository. These data include gene expression and clinical information from 50 healthy human colon mucosa along with 98 tumor CRC stage II samples and their paired adjacent morphologically normal mucosa. Raw cel files were normalized using procedure described below (see section Microarray analyses). Technical information concerning samples processing and hybridization was retrieved from the original CEL files: date of scanning were collected in order to define scan batches in each dataset separately; technical metrics described by Eklund AC and Szallasi Z in (Eklund and Szallasi, 2008) were computed and recorded as additional features for each sample. Expression data were corrected by metrics PM.IQR, RMA.IQR and RNA.DEG (Eklund and Szallasi, 2008) and by scanning day. For doing so, a linear model was fitted separately for each probeset that included these metrics as the only explanatory variables, and the coefficients of such models were used to correct the expression values a-priori. Next, the same procedure was carried out for correcting by technical effects captured by scanning date. Finally, expression intensities were summarized at the gene level (entrez) by the first principal component of the probesets mapping to the same gene. This component was centered and scaled to the weighted mean of the probesets' means and standard deviations, where the contributions to this first component were used as weights. The sign of this score was then corrected sot that it was congruent to the sign of the probeset contributing the most to the first component. Association between POL1RA expression and sample type was evaluated using a linear model (Wald test). The threshold for statistical significance was set at 5%. All analyses were carried out using R and Bioconductor (Gentleman et al., 2004).

Microarray analyses

Samples were hybridized in PrimeView Human Gene Expression Array and processed with packages affy (Gautier et al., 2004) and affyPLM (Bolstad et al., 2005) from R and Bioconductor (Gentleman et al., 2004). Raw cel files were normalized using RMA background correction and summarization (Irizarry et al., 2003). Probeset annotation was performed using the information available on the Affymetrix – Thermofisher web page (https://www.thermofisher.com/es/en/home/life-science/microarray-analysis/affymetrix.htm. Accessed on 29/10/2019). Standard quality controls were performed in order to identify abnormal samples regarding: a) spatial artifacts in the hybridization process (scan images and pseudo-images from probe level models); b) intensity dependences of differences between chips (MvA

plots); c) RNA quality (RNA digest plot); d) global intensity levels (boxplot of perfect match log-intensity distributions before and after normalization and RLE plots); e) anomalous intensity profile compared to the rest of the samples (NUSE plots, Principal Component Analyses); f) impact of quality metrics (Eklund and Szallasi, 2008) on expression measures. Samples from batch “cm.1509” in the tumor dataset (refers to the Ephb2 populations, Figure 1) were a priori corrected gene-wise by RMA.IQR metric (Eklund and Szallasi, 2008) using a linear model with no more covariates included in it. No samples were excluded from the study due to quality issues. Differential expression analysis was performed using a linear model with empirical shrinkage (Smyth, 2004) as implemented in the limma R package (Ritchie et al., 2015). Comparisons were controlled for biological replicate by including the sample’s donors in the model as covariates. In the tumor dataset (refers to the Ephb2 populations, Figure1), the models additionally included the tissue as covariate (Normal/Tumor) as well as the interaction between tissue and Ephb2 status. Benjamini and Hochberg False Discovery Rate (FDR) (Smyth, 2004) was used for multiple comparisons correction. Pathway enrichment analyses were performed using a modification of ROAST (Wu et al., 2010), a rotation-based approach implemented in the R package limma (Ritchie et al., 2015) that is specially suitable for small size experiments. Such modifications were implemented to accommodate the re-standardized maxmean statistic proposed in the ROAST algorithm, in order to enable it for competitive testing (Goeman and Bühlmann, 2007). For visualization purposes, we represented these results using the Enrichment Score graphic from the Gene Set Enrichment Analysis (GSEA) algorithm (Subramanian et al., 2005). Genesets derived from the KEGG pathway database (Kanehisa and Goto, 2000) and the Gene Ontology (GO) (Ashburner et al., 2000) as collected in the R package org.Hs.eg.db were used for these analyses, as well as the Broad Hallmark gene set collection from the MsigDB (Liberzon et al., 2015).

RNA-Seq data

Reads were aligned to the primary assembly of the mm10 Ensembl genome (GRCm38) using STAR (version 2.7.0a, default parameters) (Dobin et al., 2013). Quantification of count per genomic feature was performed using the R package Rsubread (Liao et al., 2013) (function featureCounts, default parameters). GO (Ashburner et al., 2000) terms were retrieved from R package org.Mm.eg.db , while Broad Hallmark sets were translated to mouse homologous genes using the R package biomaRt (Durinck et al., 2009). Expression counts were rlog-transformed (Love et al., 2014) and summarized at the gene set levels as follows: rlog-values

were centered and scaled gene-wise to produce z-scores, which were then averaged across all genes included in a given gene set; the resulting score were in turn centered and scaled across samples that were included in the dataset. Expression data at the gene set level were graphically represented in a heatmap using a blue to red color gradation, where red indicated the highest expression and white corresponded to the lowest expression values. For clarity, the most extreme values were truncated to -1.5 and 1.5.

Single-cell RNA-seq data analysis

Sequencing was carried out as paired-end (PE75) reads with library indexes corresponding to cell barcodes. After sequencing, libraries were inspected with the FastQC suite to assess the quality of the reads. Reads were then demultiplexed according to the cell barcodes and mapped on the human reference genome (Gencode release 27, assembly GRCh38 p10) with the RNA pipeline of the GEMTools (Marco-sola et al., 2012) 1.7.0 suite using default parameters (6% of mismatches, minimum of 80% matched bases, and minimum quality threshold of > 26).

For both PDO7 and PDO-p18, cells with <65% of mapped reads or <100,000 of total mapped reads were discarded. Cells with >100 genes and <25% of mitochondrial gene counts were included in downstream analyses, resulting in read count matrices containing 511 (PD07) and 388 (PD18) single cells. Genes that were expressed in fewer than five cells were removed. Filtering, normalization, selection of highly variable genes (HVG) and clustering of cells were performed according to the Seurat (Satija et al., 2015) package (version 2.3.4). Selection of HVG was based on the evaluation of the relationship between gene dispersion (y.cutoff = 0.5) and the log mean expression (x.high.cutoff = 3), while the total number was limited to the mean of genes per cell. Projections of HVG onto a reduced dimensional space was used for the graph-based clustering to group cells into subpopulations. At this step, the dimension of the subspace was set to the number of significant principal components (PC); based on the distribution of the PC standard deviations. The number of clusters was aligned to the expected biological variability and cluster identities were assigned using previously described gene markers and transcriptional signatures derived from bulk RNA sequencing of the principal cell populations. T-SNE was used to visualize the clustering distribution of cells. Trajectory analysis and pseudo-ordering of cells were performed with the Monocle (Qiu et al., 2017) package (version 2.8.0) using the previously identified HVG. Cell clustering and pseudo-time were consistent between and within samples.

Single-cell RNA-Seq data – enrichment analysis

LGR5/POL1 gene expression signatures were retrieved from the scRNA-Seq differential expression analyses using different thresholds for fold-change and statistical significance. These signatures were evaluated for pathway enrichment using a hypergeometric test for their intersection with gene sets in the GO (Ashburner et al., 2000) and Hallmarks (Liberzon et al., 2015) collections.

Statistical analyses

To test statistical significance between samples from two different groups two-tailed Student's *t*-tests were used. When comparing samples from the same animal paired *t*-tests were applied. When comparing different groups within different variables, multiple comparison two-way ANOVA Tukey's tests were used. For Figure 1G, 4K and 4M a mixed effects linear model was fitted with technical replicates taken as random effects and PDOs as fixed effects. The "lmer" function from the lme4 (Bates et al., 2014) R package was used for fitting the model. Coefficients and p-values for the comparisons of interest were computed through the "glht" function from the multcomp (Hothorn et al., 2008) R package using the "Westfall" multiplicity adjustment method. Confidence intervals for the coefficients were calculated with the "confint" function. Plots were generated using R programming language (Bunn, 2014). Data from 5.8S and KRT20 staining of FFPE primary CRCs (Figure 3J and S3D) were represented in a smoothed scatterplot using the R function *smoothScatter* (Bunn, 2014), where the color intensity corresponded to the density of data points through a 2-dimension kernel estimate. Joint distribution of KRT20 and 5.8S values were tested against a proper set of negative control samples stained with 5.8S and EPHB2 (data not shown). Log2-ratios of the corresponding marker against 5.8S values were computed for each stain data point. Each sample was then summarized by the median of the absolute value of these log2-ratios. Distributions of these values for KRT20 and EPHB2 samples were then compared using a Komolgorov-Smirnov test. For Kapplan-Meier curves, Gehan-Breslow-Wilcoxon test was applied. In growth kinetics experiments statistical significance was obtained by comparing the relative tumor volume of the last measure from control subjects with the corresponding measure of treated animals in a two-tailed Student's *t*-test. All group data are represented by the mean and errors bars are the standard error of the mean (s.e.m) and statistical tests and plots were generated with GraphPad Prism unless indicated.

Sample sizing and collection

All samples and animals were assigned randomly to experimental conditions, as well as the sample collection. Automated quantifications and blind data analysis were done whenever possible. A minimum of three representative images was quantified in each experiment and each condition.

EXCEL TABLE TITLE AND LEGENDS

Table S5. **Gene Signatures used in this study, Related to Figure 5 and 7.**

Gene signatures (POLR1A high, POLRA low, and LGR5) used for the scRNA-seq analysis and (Stemness, Proliferation and Differentiation) used for the RNA-seq analysis of AKP organoids.

Table S6. Gene Set Enrichment Analysis (GSEA), Related to Figure 1, 7 and S4.

GO categories enriched in EPHB2 high vs low, LGR5+ vs LGR5- in POLR1A high, and early vs long differentiation of AKP organoids.

Table S7. **Gene Set Enrichment Analysis (GSEA) of scRNA-seq, Related to Figure 5.**

GO categories enriched in POLR1A-high vs LGR5+ (PDO7) and POLR1A-high/LGR5+ vs POLR1A-low/LGR5+ (PDO-p18) populations of the scRNA-seq analysis.

REFERENCES

Ajore, R., Raiser, D., McConkey, M., Jöud, M., Boidol, B., Mar, B., Saksena, G., Weinstock, D.M., Armstrong, S., Ellis, S.R., et al. (2017). Deletion of ribosomal protein genes is a common vulnerability in human cancer, especially in concert with TP53 mutations. EMBO Mol. Med. *9*, 498–507.

Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P., Dolinski, K., Dwight, S.S., Eppig, J.T., et al. (2000). Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat. Genet. *25*, 25–29.

Aytes, A., Molleví, D.G., Martinez-Iniesta, M., Nadal, M., Vidal, A., Morales, A., Salazar, R., Capellà, G., and Villanueva, A. (2012). Stromal interaction molecule 2 (STIM2) is frequently overexpressed in colorectal tumors and confers a tumor cell growth suppressor phenotype. Mol. Carcinog. *51*, 746–753.

Barna, M., Pusic, A., Zollo, O., Costa, M., Kondrashov, N., Rego, E., Rao, P.H., and Ruggero, D. (2008). Suppression of Myc oncogenic activity by ribosomal protein haploinsufficiency. Nature *456*, 971–975.

Bates, D., Mächler, M., Bolker, B., and Walker, S. (2014). Fitting Linear Mixed-Effects Models using lme4.

Batlle, E., Henderson, J.T., Beghtel, H., Born, M.M.W. Van Den, Sancho, E., Huls, G., Meeldijk, J., Robertson, J., Wetering, M. Van De, Pawson, T., et al. (2002). b-Catenin and TCF Mediate Cell Positioning in the Intestinal Epithelium by Controlling the Expression of EphB / EphrinB. Cell *111*, 251–263.

Blanco, S., Bandiera, R., Popis, M., Hussain, S., Lombard, P., Aleksic, J., Sajini, A., Tanna, H., Cortés-Garrido, R., Gkatza, N., et al. (2016). Stem cell function and stress response are controlled by protein synthesis. Nature *534*, 335–340.

Bolstad, B.M., Collin, F., Brettschneider, J., Simpson, K., Cope, L., Irizarry, R.A., and Speed, T.P. (2005). Quality Assessment of Affymetrix GeneChip Data. In Bioinformatics and Computational Biology Solutions Using R and Bioconductor, (New York: Springer-Verlag), pp. 33–47.

Bunn, A. (2014). Crossdating in dplR. *2*, 1–12.

Bustelo, X.R., and Dosil, M. (2018). Ribosome biogenesis and cancer: basic and translational challenges. Curr. Opin. Genet. Dev. *48*, 22–29.

Calon, A., Espinet, E., Palomo-Ponce, S., Tauriello, D.V.F., Iglesias, M., Céspedes, M.V., Sevillano, M., Nadal, C., Jung, P., Zhang, X.H.F., et al. (2012). Dependency of Colorectal Cancer on a TGF-β-Driven Program in Stromal Cells for Metastasis Initiation. Cancer Cell *22*, 571–584.

Calve, S., Witten, A.J., Ocken, A.R., and Kinzer-Ursem, T.L. (2016). Incorporation of non-canonical amino acids into the developing murine proteome. Sci. Rep. *6*, 32377.

Cortina, C., Turon, G., Stork, D., Hernando□Momblona, X., Sevillano, M., Aguilera, M., Tosi, S., Merlos□Suárez, A., Stephan□Otto Attolini, C., Sancho, E., et al. (2017). A genome editing approach to study cancer stem cells in human tumors. EMBO Mol. Med. *9*, 869–879.

Dalerba, P., Dylla, S.J., Park, I.K., Liu, R., Wang, X., Cho, R.W., Hoey, T., Gurney, A., Huang, E.H., Simeone, D.M., et al. (2007). Phenotypic characterization of human colorectal cancer stem cells. Proc.Natl.Acad.Sci.U.S.A *104*, 10158–10163.

Dalerba, P., Kalisky, T., Sahoo, D., Rajendran, P.S., Rothenberg, M.E., Leyrat, A.A., Sim, S., Okamoto, J., Johnston, D.M., Qian, D., et al. (2011). Single-cell dissection of transcriptional heterogeneity in human colon tumors. Nat. Biotechnol. *29*, 1120–1127.

Dieterich, D.C., Hodas, J.J.L., Gouzer, G., Shadrin, I.Y., Ngo, J.T., Triller, A., Tirrell, D.A., and Schuman, E.M. (2010). In situ visualization and dynamics of newly synthesized proteins in rat hippocampal neurons. Nat. Neurosci. *13*, 897–905.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: Ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15–21.

Dow, L.E., O'Rourke, K.P., Simon, J., Tschaharganeh, D.F., van Es, J.H., Clevers, H., and Lowe, S.W. (2015). Apc Restoration Promotes Cellular Differentiation and

Reestablishes Crypt Homeostasis in Colorectal Cancer. Cell *161*, 1539–1552.

Durinck, S., Spellman, P.T., Birney, E., and Huber, W. (2009). Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. Nat. Protoc. *4*, 1184–1191.

Eklund, A.C., and Szallasi, Z. (2008). Correction of technical bias in clinical microarray data improves concordance with known biological information. Genome Biol. *9*, R26.

Fumagalli, A., Oost, K.C., Kester, L., Morgner, J., Bornes, L., Bruens, L., Spaargaren, L., Azkanaz, M., Schelfhorst, T., Beerling, E., et al. (2020). Plasticity of Lgr5-Negative Cancer Cells Drives Metastasis in Colorectal Cancer. Cell Stem Cell.

Gautier, L., Cope, L., Bolstad, B.M., and Irizarry, R.A. (2004). affy--analysis of Affymetrix GeneChip data at the probe level. Bioinformatics *20*, 307–315.

Gentleman, R.C., Carey, V.J., Bates, D.M., Bolstad, B., Dettling, M., Dudoit, S., Ellis, B., Gautier, L., Ge, Y., Gentry, J., et al. (2004). Bioconductor: open software development for computational biology and bioinformatics. Genome Biol. *5*, R80.

Goeman, J.J., and Bühlmann, P. (2007). Analyzing gene expression data in terms of gene sets: Methodological issues. Bioinformatics *23*, 980–987.

Gonzalez-Roca, E., Garcia-Albéniz, X., Rodriguez-Mulero, S., Gomis, R.R., Kornacker, K., and Auer, H. (2010). Accurate expression profiling of very small cell populations. PLoS One *5*, e14418.

van der Heijden, M., Miedema, D.M., Waclaw, B., Veenstra, V.L., Lecca, M.C., Nijman, L.E., van Dijk, E., van Neerven, S.M., Lodestijn, S.C., Lenos, K.J., et al. (2019). Spatiotemporal regulation of clonogenicity in colorectal cancer xenografts. Proc. Natl. Acad. Sci. *116*, 6140–6145.

Hothorn, T., Bretz, F., and Westfall, P. (2008). Simultaneous Inference in General Parametric Models. *50*, 346–363.

Irizarry, R.A., Hobbs, B., Beazer-barclay, Y.D., Antonellis, K.J., Scherf, U.W.E., and

Speed, T.P. (2003). Exploration , normalization , and summaries of high density oligonucleotide array probe level data. 249–264.

Jao, C.Y., and Salic, A. (2008). Exploring RNA transcription and turnover in vivo by using click chemistry. Proc. Natl. Acad. Sci. U. S. A. *105*, 15779–15784.

Jiang, X., Hao, H.-X., Growney, J.D., Woolfenden, S., Bottiglio, C., Ng, N., Lu, B., Hsieh, M.H., Bagdasarian, L., Meyer, R., et al. (2013). Inactivating mutations of RNF43 confer Wnt dependency in pancreatic ductal adenocarcinoma. Proc. Natl. Acad. Sci. *110*, 12649–12654.

Kanehisa, M., and Goto, S. (2000). KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res. *28*, 27–30.

Kozar, S., Morrissey, E., Nicholson, A.M., van der Heijden, M., Zecchini, H.I., Kemp, R., Tavaré, S., Vermeulen, L., and Winton, D.J. (2013). Continuous Clonal Labeling Reveals Small Numbers of Functional Stem Cells in Intestinal Crypts and Adenomas. Cell Stem Cell *13*, 626–633.

Lenos, K.J., Miedema, D.M., Lodestijn, S.C., Nijman, L.E., van den Bosch, T., Romero Ros, X., Lourenço, F.C., Lecca, M.C., van der Heijden, M., van Neerven, S.M., et al. (2018). Stem cell functionality is microenvironmentally defined during tumour expansion and therapy response in colon cancer. Nat. Cell Biol. *20*, 1193–1202.

Lerner, E.A., Lerner, M.R., Janeway, C.A., and Steitz, J.A. (1981). Monoclonal antibodies to nucleic acid-containing cellular constituents: probes for molecular biology and autoimmune disease. Proc. Natl. Acad. Sci. U. S. A. *78*, 2737–2741.

Liao, Y., Smyth, G.K., and Shi, W. (2013). The Subread aligner: fast, accurate and scalable read mapping by seed-and-vote. Nucleic Acids Res. *41*, e108.

Liberzon, A., Birger, C., Thorvaldsdóttir, H., Ghandi, M., Mesirov, J.P., and Tamayo, P. (2015). The Molecular Signatures Database Hallmark Gene Set Collection. Cell Syst. *1*, 417–425.

Liu, J., Xu, Y., Stoleru, D., and Salic, A. (2012). Imaging protein synthesis in cells

and tissues with an alkyne analog of puromycin. Proc. Natl. Acad. Sci. *109*, 413–418.

Lombardo, Y., Scopelliti, A., Cammareri, P., Todaro, M., Iovino, F., Ricci-Vitiani, L., Gulotta, G., Dieli, F., De Maria, R., and Stassi, G. (2011). Bone morphogenetic protein 4 induces differentiation of colorectal cancer stem cells and increases their response to chemotherapy in mice. Gastroenterology *140*, 297–309.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. *15*.

Madan, B., Harmston, N., Nallan, G., Montoya, A., Faull, P., Petretto, E., and Virshup, D.M. (2018). Temporal dynamics of Wnt-dependent transcriptome reveal an oncogenic Wnt/MYC/ribosome axis. J. Clin. Invest. *128*, 5620–5633.

Marco-sola, S., Sammeth, M., Guigó, R., and Ribeca, P. (2012). The GEM mapper : fast , accurate and versatile alignment by filtration. *9*.

Merlos-Suárez, A., Barriga, F.M., Jung, P., Iglesias, M., Céspedes, M.V., Rossell, D., Sevillano, M., Hernando-Momblona, X., da Silva-Diz, V., Muñoz, P., et al. (2011).

The Intestinal Stem Cell SignatureIdentifies Colorectal Cancer Stem Cells and Predicts Disease Relapse. Cell Stem Cell 1–14.

O'Brien, C.A., Pollett, A., Gallinger, S., and Dick, J.E. (2007). A human colon cancer cell capable of initiating tumour growth in immunodeficient mice. Nature *445*, 106–110.

Pelletier, J., Thomas, G., and Volarević, S. (2017). Ribosome biogenesis in cancer: new players and therapeutic avenues. Nat. Rev. Cancer *18*, 51–63.

Peltonen, K., Colis, L., Liu, H., Trivedi, R., Moubarek, M.S., Moore, H.M., Bai, B., Rudek, M.A., Bieberich, C.J., and Laiho, M. (2014). A Targeting Modality for Destruction of RNA Polymerase I that Possesses Anticancer Activity. Cancer Cell *25*, 77–90.

Picelli, S., Björklund, Å.K., Faridani, O.R., Sagasser, S., Winberg, G., and Sandberg, R. Smart-seq2 for sensitive full-length transcriptome profiling in single cells.

Qiu, X., Hill, A., Packer, J., Lin, D., Ma, Y., and Trapnell, C. (2017). Single-cell mRNA quantification and differential analysis with Census.

Ricci-Vitiani, L., Lombardi, D.G., Pilozzi, E., Biffoni, M., Todaro, M., Peschle, C., and De, M.R. (2007). Identification and expansion of human colon-cancer-initiating cells. Nature *445*, 111–115.

Ritchie, M.E., Phipson, B., Wu, D., Hu, Y., Law, C.W., Shi, W., and Smyth, G.K. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. *43*.

Robichaud, N., and Sonenberg, N. (2017). Translational control and the cancer cell response to stress. Curr. Opin. Cell Biol. *45*, 102–109.

Robichaud, N., Sonenberg, N., Ruggero, D., and Schneider, R.J. (2019).

Translational Control in Cancer. Cold Spring Harb. Perspect. Biol. *11*, a032896.
Ruggero, D. (2012). Revisiting the Nucleolus: From Marker to Dynamic Integrator of Cancer Signaling. Sci. Signal. *5*, pe38-pe38.

Sanz-Pamplona, R., Berenguer, A., Cordero, D., Molleví, D.G., Crous-Bou, M., Sole, X., Paré-Brunet, L., Guino, E., Salazar, R., Santos, C., et al. (2014). Aberrant gene expression in mucosa adjacent to tumor reveals a molecular crosstalk in colon cancer. Mol. Cancer *13*, 46.

Satija, R., Farrell, J.A., Gennert, D., Schier, A.F., and Regev, A. (2015). Spatial reconstruction of single-cell gene expression data. Nat Biotechnol *33*.

Schindelin, J., Arganda-Carrera, I., Frise, E., Verena, K., Mark, L., Tobias, P., Stephan, P., Curtis, R., Stephan, S., Benjamin, S., et al. (2009). Fiji - an Open platform for biological image analysis. Nat. Methods *9*.

Shimokawa, M., Ohta, Y., Nishikori, S., Matano, M., Takano, A., Fujii, M., Date, S., Sugimoto, S., Kanai, T., and Sato, T. (2017). Visualization and targeting of LGR5 + human colon cancer stem cells. Nature *545*, 187–192.

Signer, R.A.J., Magee, J.A., Salic, A., and Morrison, S.J. (2014). Haematopoietic stem cells require a highly regulated protein synthesis rate. Nature *509*, 49–54.

Smyth, G.K. (2004). Linear Models and Empirical Bayes Methods for Assessing Differential Expression in Microarray Experiments. Stat. Appl. Genet. Mol. Biol. *3*, 1–25.

de Sousa e Melo, F., Kurtova, A. V., Harnoss, J.M., Kljavin, N., Hoeck, J.D., Hung, J., Anderson, J.E., Storm, E.E., Modrusan, Z., Koeppen, H., et al. (2017). A distinct role for Lgr5+ stem cells in primary and metastatic colon cancer. Nature *543*, 676–680.

Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., et al. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc. Natl. Acad. Sci. U. S. A. *102*, 15545–15550.

Tauriello, F., Palomo-ponce, S., Stork, D., Berenguer-llergo, A., Badia-ramentol, J., Iglesias, M., Sevillano, M., Ibiza, S., Cañellas, A., Hernando-momblona, X., et al. (2018). TGF β drives immune evasion in genetically. Nat. Publ. Gr. *554*, 538–543.

van de Wetering, M., Francies, H.E., Francis, J.M., Bounova, G., Iorio, F., Pronk, A., van Houdt, W., van Gorp, J., Taylor-Weiner, A., Kester, L., et al. (2015). Prospective Derivation of a Living Organoid Biobank of Colorectal Cancer Patients. Cell *161*, 933–945.

Vermeulen, L., Todaro, M., de Sousa Mello, F., Sprick, M.R., Kemper, K., Perez Alea, M., Richel, D.J., Stassi, G., and Medema, J.P. (2008). Single-cell cloning of colon cancer stem cells reveals a multi-lineage differentiation capacity. Proc Natl Acad Sci U S A *105*, 13427–13432.

Vermeulen, L., De Sousa, E.M.F., van der Heijden, M., Cameron, K., de Jong, J.H., Borovski, T., Tuynman, J.B., Todaro, M., Merz, C., Rodermond, H., et al. (2010). Wnt activity defines colon cancer stem cells and is regulated by the microenvironment. Nat Cell Biol *12*, 468–476.

van de Wetering, M., Sancho, E., Verweij, C., de, L.W., Oving, I., Hurlstone, A., van

der, H.K., Batlle, E., Coudreuse, D., Haramis, A.P., et al. (2002). The beta-catenin/TCF-4 complex imposes a crypt progenitor phenotype on colorectal cancer cells. Cell *111*, 241–250.

Witte, D.P., Wiginton, D.A., Hutton, J.J., and Aronow, B.J. (1991). Coordinate developmental regulation of purine catabolic enzyme expression in gastrointestinal and postimplantation reproductive tracts. J Cell Biol *115*, 179–190.

Wu, D., Lim, E., Vaillant, F., Asselin-Labat, M.-L., Visvader, J.E., and Smyth, G.K. (2010). ROAST: rotation gene set tests for complex microarray experiments. Bioinformatics *26*, 2176–2182.

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Goat Anti Mouse EphB2	R&D systems	Cat# AF467 RRID: AB_355375
Goat Anti Human EpCAM /TROP-1	R&D systems	Cat# AF960 RRID: 355745
Anti-Ki67	BD Biosciences	Cat# 550609 RRID: 393778
RPA194 (C-1) antibody	Santa Cruz	Cat# sc-48385 RRID: AB_675814
tdTomato polyclonal antibody	Sicgen	Cat# AB8181-200 RRID: AB_2722750
Anti-RFP/tdTomato antibody	Rockland	Cat# 600-401-379 RRID: AB_2209751
GFP antibody	Abcam	Cat# ab183734 RRID: AB_2732027
Cytokeratin 20 antibody	Dako	Cat# M7019 RRID: AB_2133718
Anti-Mouse CD16 / CD32	Tonbo Biosciences	Cat# 70-0161 RRID: AB_2621487
BV421 Rat Anti-Mouse CD31 Clone MEC 13.3	BD Biosciences	Cat# 562939 RRID: AB_2665476
BV421 Mouse CD45RB Clone 16A	BD Biosciences	Cat# 562849 clone16A RRID: AB_2737836

CD326 (EpCAM) Monoclonal Antibody (1B7), PE-Cyanine7	Thermo Fisher Scientific	Cat# 25-9326-42 RRID: AB_2573542
CD326 (EpCAM)-APC-Vio770 human antibody	Miltenyl Biotec	Cat# 130-101-161 RRID: AB_2660308
APC/Cyanine7 anti-mouse CD326 (Ep-CAM) antibody	Biolegends	Cat# 118217, RRID: AB_1501158
APC anti-human EphB2 antibody (2H9 clone)	BD Biosciences	Cat# 564699 RRID: AB_2738898
rRNA(5.8S)(Y10b)	Thermo Fisher Scientific	Cat# MA1-13017, RRID: AB_10979967
Donkey anti-Goat conjugated to Alexa-488/568/647	Thermo Fisher Scientific	Cat# A11055 RRID: AB_2534102 Cat# A11057 RRID: AB_142581 Cat# A21447 RRID: AB_141844
Donkey anti-Rabbit conjugated to Alexa-488/568/647	Thermo Fisher Scientific	Cat# A21206 RRID: AB_2535792 Cat# A10042 RRID: AB_2534017 Cat# A31573 RRID: AB_2536183

Donkey anti-Mouse conjugated to Alexa-488/568/647	Thermo Fisher Scientific	Cat# A21202 RRID: 141607 Cat# A10037 RRID: 2534013 Cat# A31571 RRID: AB_162542
BightVision poly-HRP anti-Rabbit	Immunologic	DPVR-110HRP
Rabbit anti-RRN3	Sigma-Aldrich/Merck	HPA049837
Mouse anti-bActin	Abcam	20272
KRT20	Atlas Antibodies	HPA024309
Goat anti-Rabbit Alexa Fluor 680	Thermo Fisher Scientific	A-21076
Biological samples		
Patient derived organoids (PDOs	(van de Wetering et al., 2015) Giorgio Stassi Laboratory (Lombardo et al., 2011)	n/a
Mouse derived organoids (shAPC-KP)	(Dow et al., 2015)	n/a
Experimental models: cell lines		
LS174T	ATCC	CL-188
SW403	ATCC	CCL-230
Chemicals, Peptides, and Recombinant Proteins		
Tamoxifen	Sigma-Aldrich/Merck	T5648
4-hydroxytamoxifen	Peprotech Sigma-Aldrich/Merck	6833585 H7904
Doxycicline hyclate	Sigma-Aldrich/Merck	D9891
5-Fluorouracil (5-FU)	Sigma-Aldrich/Merck	F6627

Irinotecan	Sigma-Aldrich/Merck	I1406
Leucovorin	Sigma-Aldrich/Merck	47612
TrypLE™ Express Enzyme (1X)	Thermo Fisher Scientific	12604039
Trypsine	Thermo Fisher Scientific	25300054
DMEM high glucose pyruvate	Thermo Fisher Scientific	41966052
Advanced-DMEM/F-12	Thermo Fisher Scientific	12634028
Matrigel growth factor reduced	Corning	356231
Cultrex Growth Factor Reduced BME2	Amsbio	3533-010-02
HEPES	Thermo Fisher Scientific	15630056
Glutamax	Thermo Fisher Scientific	35050-038
B-27 Supplement	Thermo Fisher Scientific	12587-010
N-Acetyl-L-cysteine	Sigma-Aldrich/Merck	A7250-
Nicotinamide	Sigma-Aldrich/Merck	72340
EGF	Peprotech	AF-100-15
bFGF	Thermo Fisher Scientific	PHG0026
Noggin	In house	n/a
R-spondin	In house	n/a
Rock inhibitor, Y-27632	Med Chem express	HY-10583
Gastrin I	Tocris	3006
TGFb-RI inhibitor, A83-01	Tocris	2939

P38a/b inhibitor, SB202190 monohydrochloride hydrate	Sigma-Aldrich/Merck	S7076
TGFb-RI inhibitor (galunisertib), LY2157299	In house	n/a
Chemical Inducer of Dimerization, AP20187	Med Chem express	HY-13992
5-Ethynyl-2'-deoxyuridine (EdU)	Thermo Fisher Scientific	A10044
5-Ethynyl uridine (EU)	Click chemistry Tools	1261-100
O-propargyl-puromycin (OPP)	Medchem Source	JA-1024
L-homopropargylglycine (HPG)	Click chemistry Tools	1067-100
Polymerase I inhibitor, BMH-21	Invivochem	V1435
Cell recovery solution	Corning	354253
Proteinase K	Thermo Fisher Scientific	25530049
IRDye 800CW Azide Infrared Dye	Li-COR (Bonsai Advanced Technologies)	929-60000
Smart-seq2 lysis buffer	In house Picelli S. et al., 2014	n/a
Critical commercial assays		
High-Capacity cDNA Reverse Transcription kit	Thermo Fisher Scientific	4368813
Nucleospin Plasmid QuickPure Miniprep Kit	Cultek	740615250
PureLink® RNA Mini Kit	Thermo Fisher Scientific	12183025
GenElute Mammalian Genomic DNA Miniprep Kit	Sigma-Aldrich/Merck	G1N70-1KT
Cell Line Nucleofector® Kit V	Cultek (Lonza)	VCA-1003
Click-iT™ Edu Alexa Fluor 647 Imaging Kit	Thermo Fisher Scientific	C10340
Click-iT™ RNA Alexa Fluor 488 Imaging Kit	Thermo Fisher Scientific	C10329

Click-iT™ Plus OPP Alexa Fluor™ 594 Protein Synthesis Assay Kit	Thermo Fisher Scientific	C10457
SuperSignal West Pico Chemiluminescent Substrate	Thermo Fisher Scientific	34080
GeneChip Human PrimeView array	Affymetrix	901837
GeneChip Mapping 250 K Nsp assay kit	Affymetrix	900753
GeneChip Hybridization Wash and Stain Kit	Affymetrix	900720
GeneChip Fluidics Station 450	Affymetrix	00-0377
High sensitivity DNA Kit	Agilent Techonologies	5067-4626
Nextera® XT DNA Library Prep Kit ® XT	Illumina	FC-131-1024
MegaPrime labeling kit	GE Healthcare	RPN1604

Reagent or resource		
SYBR® Green PCR Master Mix	Thermo Fisher Scientific	4368706
TaqMan® Universal PCR Master Mix	Thermo Fisher Scientific	4364341
Phusion High Fidelity DNA Polymerase	Thermo Fisher Scientific	F530L
DNA Polymerase	Biotools	10.014
SuperScript II Reverse Transcriptase	Thermo Fisher Scientific	18064022
Agencourt Ampure XP beads	Beckman coulter	A63880
KAPA HiFi Hotstart ReadyMix	Roche	KR0370
Trizol	Thermo Fisher Scientific	15596018
Schiff's Reagent	Panreac	171588.1609
Alcian Blue solution	Panreac	CA254584.1604
DPX	Panreac	255254
Peroxidase blocking solution	Dako	S202386
Envision Flex antibody diluent	Agilent	K8006

Normal Donkey Serum	Jackson Immunoresearch	017-000-121
Normal Goat Serum	Jackson immunoresearch	005-000-121
DAPI-Fluoromount-G Mounting Medium	Southern Biotech	0100-20
NucBlue Live ReadyProbes Reagent, Hoechst	Thermo Fisher Scientific	R37605
Oligonucleotide		
sgRNA guides	This paper	Table S1
Genotyping	This paper	Table S2
Southern probes	This paper	Table S3
Taqman primers	This paper	STAR Methods
Deposited data		
Microarrays and RNA-seq	This paper	GEO: GSE125232
Single cell RNA-seq	This paper	GEO: GSE148345
Software and Algorithms		
FlowJo	BD Biosciences	https://flowjo.com
NDP.view2	n/a	www.hamamatsu.com
Fiji	Schindelin et al., 2012	https://imagej.net/Fiji
GraphPad Prism	Prism software	RRID: SCR_002798
Illustrator CS6	Adobe	RRID: SCR_010279
StepOne 2.2 plus and SDS2.4	Thermofisher	www.thermofisher.com
BD FACS Diva Software	BD Biosciences	RRID: SCR_001456
ScanR Software	n/a	www.olimpus-lifesciences.com
Zeiss LSM 780 Software	n/a	www.zeiss.com
QuPath	n/a	https://qupath.github.io/
R language and environment for statistical computing	R core team (2019)	https://cran.r-project.org/
affy package	Gautier et al., 2004	https://bioconductor.org
affy PLM package	Bolstad et al., 2005	https://bioconductor.org
Limma R package	Ritchie et al., 2015	https://bioconductor.org

STAR version 2.7.0a	Dobin et al., 2013	https://github.com
R package Rsubread	Liao et al., 2019	https://bioconductor.org
R package BiomaRt	Durinck et al., 2009	https://bioconductor.org
DESeq2 R package	Love MI et al., 2014	https://bioconductor.org
GSEA	Subramanian et al., 2005	https://www.gsea-msigdb.org/gsea/index.jsp
FastQC Suite	n/a	https://www.bioinformatics.babraham.ac.uk
GEMT Tools 1.7.0.	Marco-Sola et al., 2012	http://gemtools.github.io
Seurat package 2.3.4	Satija et al., 2015	https://bioconductor.org
Monocle package 2.8.0	Qiu et al., 2017	https://bioconductor.org
lme4 R package	Bates et al., 2014	https://cran.r-project.org
Multcomp R package	Hothorn et al., 2008	https://cran.r-project.org

Figure 1

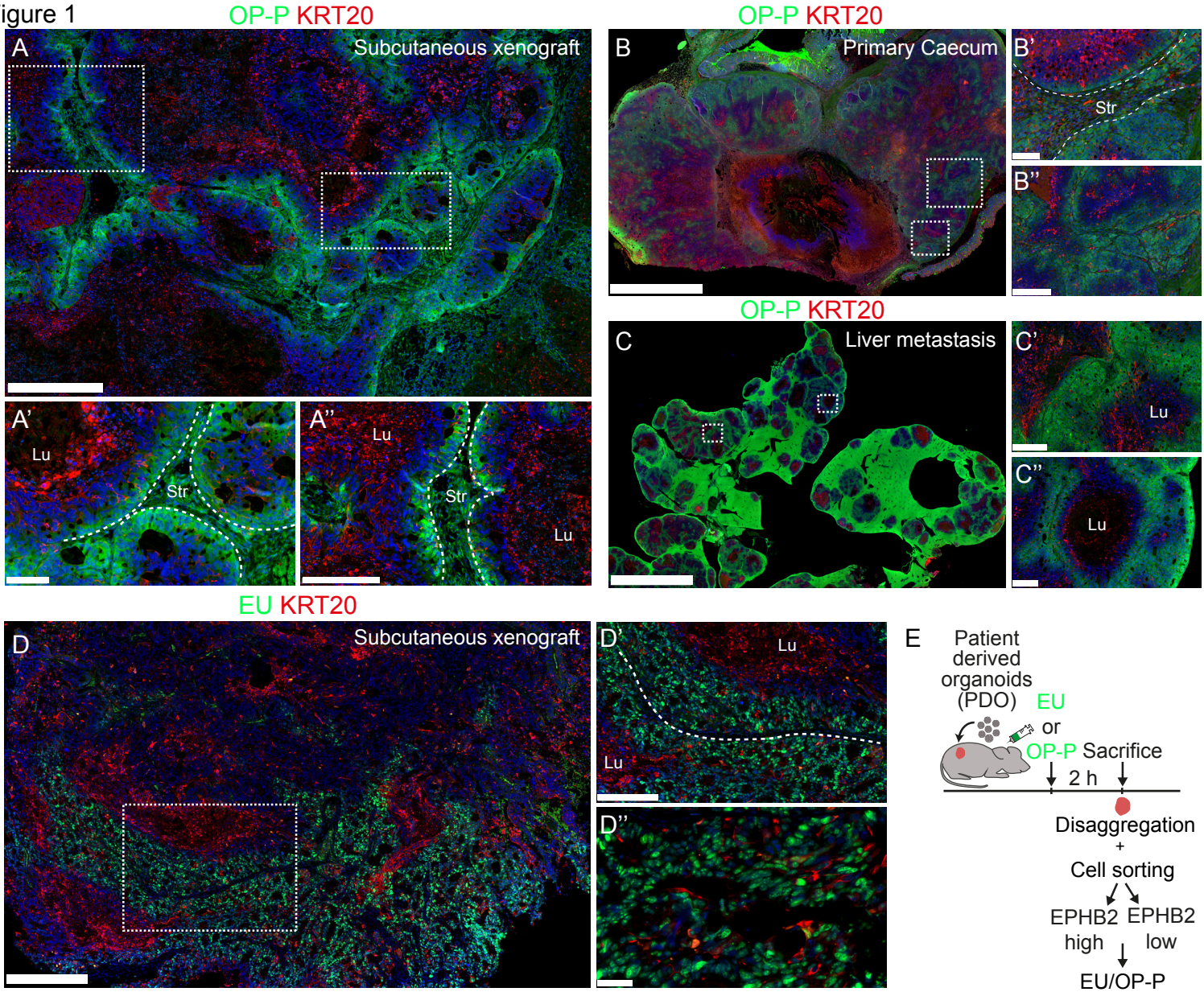

OP-P KRT20
A
Subcutaneous xenograft
A'
Lu
Str
A''
Lu
Str
Lu
OP-P KRT20
B
Primary Caecum
B'
Str
B''
OP-P KRT20
C
Liver metastasis
C'
Lu
C''
Lu
EU KRT20
D
Subcutaneous xenograft
D'
Lu
D''
E
Patient derived organoids (PDO)
EU
or
OP-P
Sacrifice
2 h
Disaggregation
+
Cell sorting
EPHB2 high
EPHB2 low
EU/OP-P
F
PDO19b
PDO-p18
PDO7
Normalized to cell count
NO OP-P
EPHB2 high
EPHB2 low
EU
OP-P
G
OP-P
EU
% of OP-P+ cells (log)
% of EU+ cells (log)
****
High
Low
EPHB2 population
Tumor: ■PDO-p18 ▲PDO19b ●PDO7
H
Tumor cell disaggregation
CRC
Cell sorting
EPHB2 high
EPHB2 low
I
EPHB2 high vs low
GOSLIM
Ribosome biogenesis
Structural constituent of ribosome
Translation
Protein folding
Nucleolus
Cellular aminoacid metabolic process
>1,5
NES
>2
J
Differentiation
OP-P
EU
Protein synthesis
rDNA transcription
Stroma
Tumor gland
Lumen

Figure 2

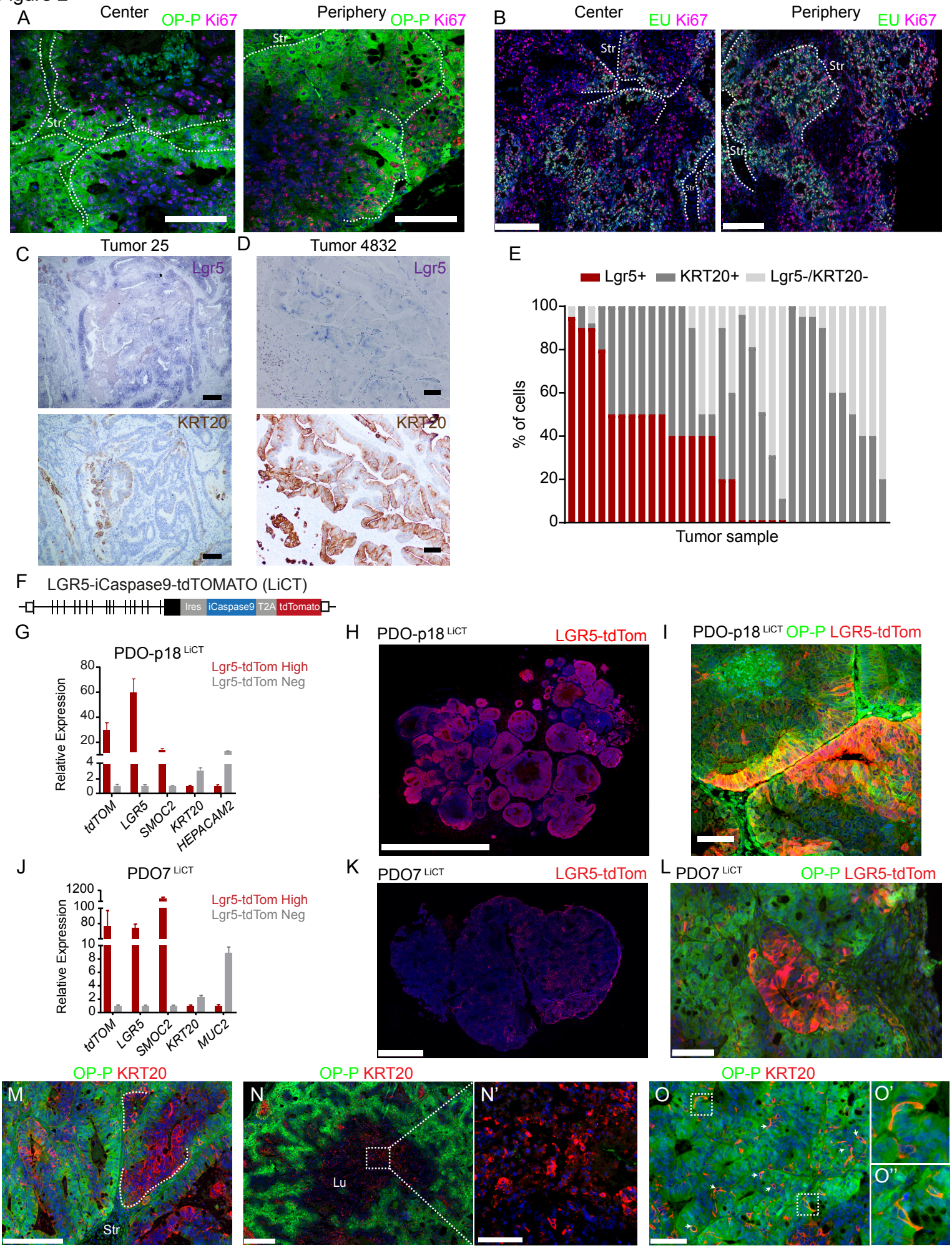

A
Center
OP-P Ki67
Periphery
OP-P Ki67
Str
B
Center
EU Ki67
Periphery
EU Ki67
C
Tumor 25
Lgr5
KRT20
D
Tumor 4832
E
Lgr5+
KRT20+
Lgr5-/KRT20-
% of cells
Tumor sample
F
LGR5-iCaspase9-tdTOMATO (LiCT)
Ires
iCaspase9
T2A
tdTomato
G
PDO-p18 LiCT
Lgr5-tdTom High
Lgr5-tdTom Neg
Relative Expression
tdTOM
LGR5
SMOC2
KRT20
HEPACAM2
H
PDO-p18 LiCT
LGR5-tdTom
I
PDO-p18 LiCT
OP-P LGR5-tdTom
J
PDO7 LiCT
Lgr5-tdTom High
Lgr5-tdTom Neg
Relative Expression
tdTOM
LGR5
SMOC2
KRT20
MUC2
K
PDO7 LiCT
LGR5-tdTom
L
PDO7 LiCT
OP-P LGR5-tdTom
OP-P KRT20
M
Str
N
Lu
N'
O
O'
O''

Figure 3

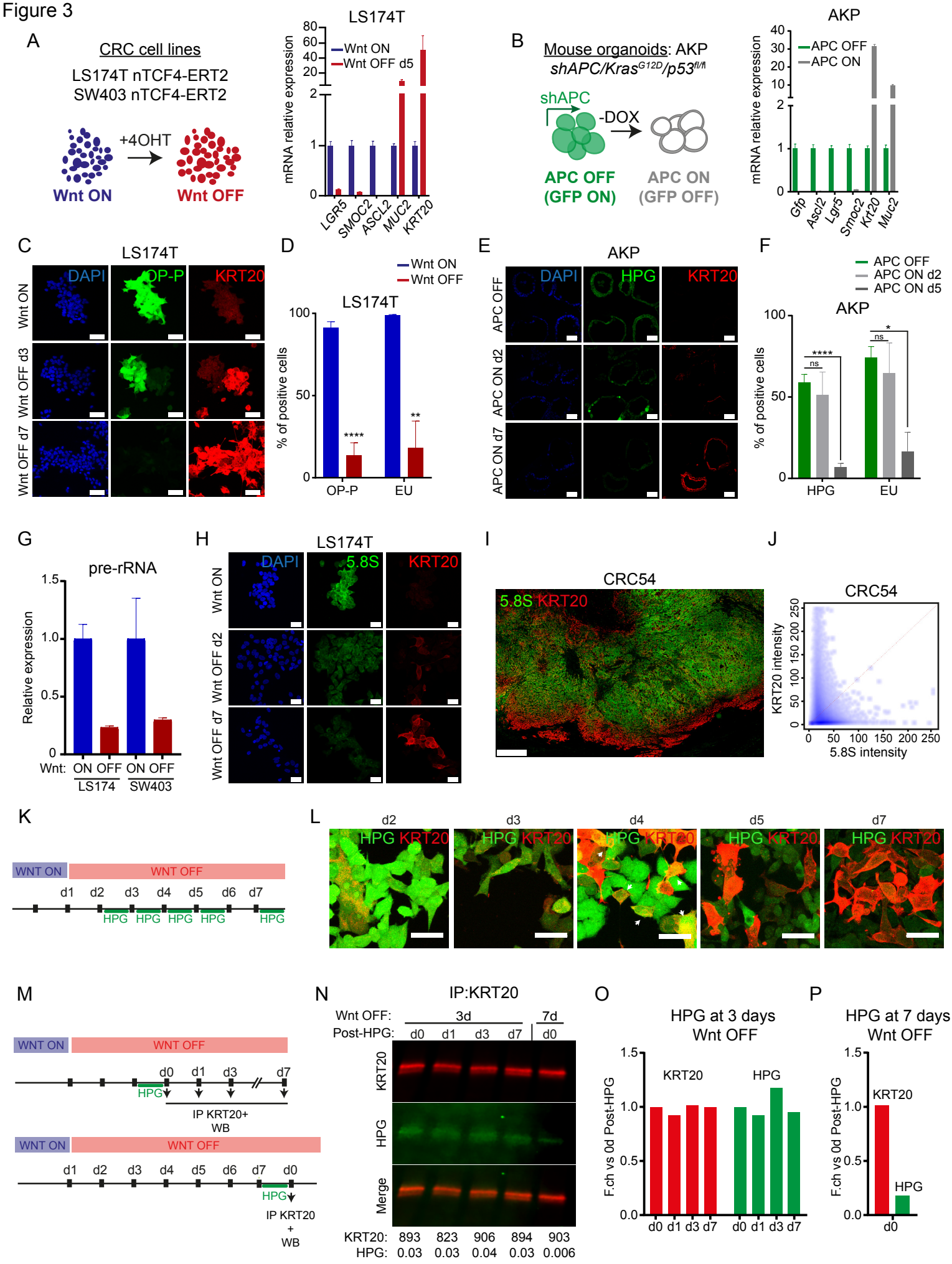

A
CRC cell lines
LS174T nTCF4-ERT2
SW403 nTCF4-ERT2
+4OHT
Wnt ON
Wnt OFF
LS174T
mRNA relative expression
Wnt ON
Wnt OFF d5
LGR5
SMOC2
ASCL2
MUC2
KRT20
B
Mouse organoids: AKP
shAPC/Kras^G12D/p53^fl/fl
shAPC
-DOX
APC OFF (GFP ON)
APC ON (GFP OFF)
AKP
APC OFF
APC ON
Gfp
Ascl2
Lgr5
Smoc2
Krt20
Muc2
C
LS174T
DAPI
OP-P
KRT20
Wnt ON
Wnt OFF d3
Wnt OFF d7
D
LS174T
Wnt ON
Wnt OFF
% of positive cells
OP-P
EU
E
AKP
DAPI
HPG
KRT20
APC OFF
APC ON d2
APC ON d7
F
AKP
APC OFF
APC ON d2
APC ON d5
% of positive cells
HPG
EU
G
pre-rRNA
Relative expression
Wnt: ON OFF ON OFF
LS174
SW403
H
LS174T
DAPI
5.8S
KRT20
Wnt ON
Wnt OFF d2
Wnt OFF d7
I
CRC54
5.8S KRT20
J
CRC54
KRT20 intensity
5.8S intensity
K
WNT ON
WNT OFF
d1 d2 d3 d4 d5 d6 d7
HPG HPG HPG HPG HPG
L
d2
d3
d4
d5
d7
HPG KRT20
M
WNT ON
WNT OFF
d0 d1 d3 d7
HPG
IP KRT20+
WB
WNT ON
WNT OFF
d1 d2 d3 d4 d5 d6 d7 d0
HPG
IP KRT20
+
WB
N
IP:KRT20
Wnt OFF: 3d 7d
Post-HPG: d0 d1 d3 d7 d0
KRT20
HPG
Merge
KRT20: 893 823 906 894 903
HPG: 0.03 0.03 0.04 0.03 0.006
O
HPG at 3 days
Wnt OFF
KRT20
HPG
F.ch vs 0d Post-HPG
d0 d1 d3 d7 d0 d1 d3 d7
P
HPG at 7 days
Wnt OFF
KRT20
HPG
F.ch vs 0d Post-HPG
d0

Figure 4

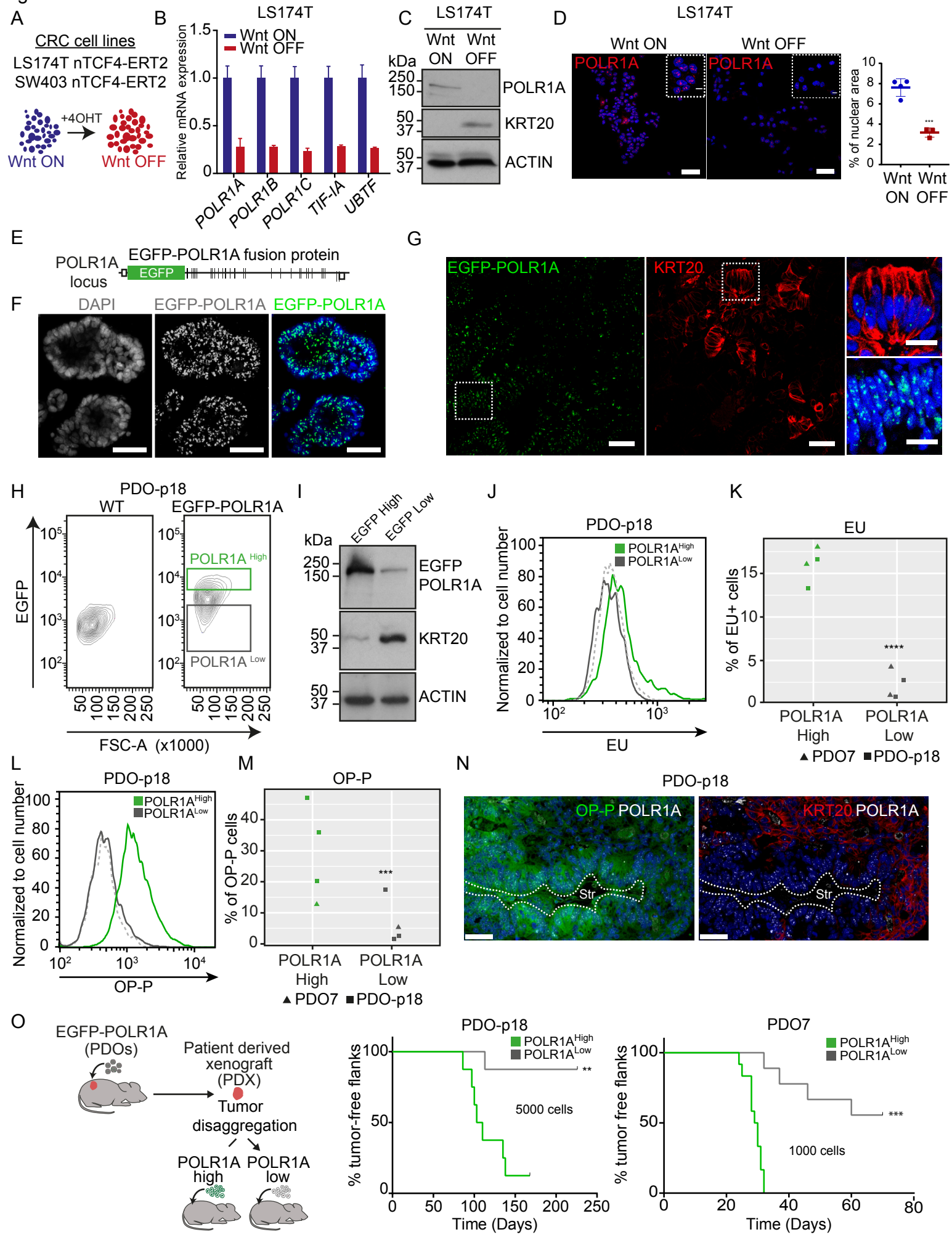
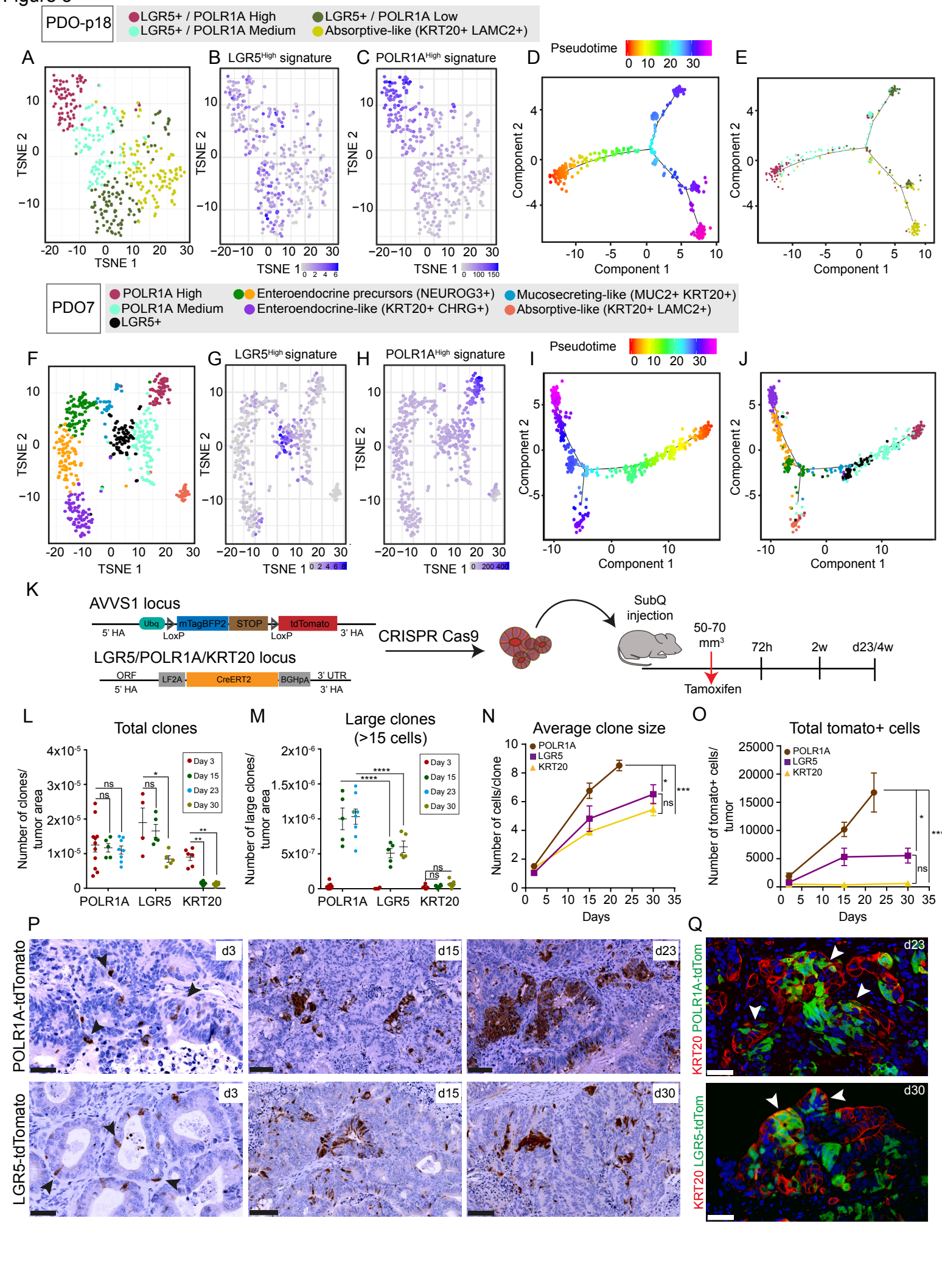

A
CRC cell lines
LS174T nTCF4-ERT2
SW403 nTCF4-ERT2
+4OHT
Wnt ON
Wnt OFF
B
LS174T
Wnt ON
Wnt OFF
Relative mRNA expression
POLR1A
POLR1B
POLR1C
TIF-IA
UBTF
C
LS174T
Wnt ON
Wnt OFF
kDa
POLR1A
KRT20
ACTIN
D
LS174T
Wnt ON
Wnt OFF
POLR1A
% of nuclear area
E
POLR1A locus
EGFP-POLR1A fusion protein
EGFP
F
DAPI
EGFP-POLR1A
EGFP-POLR1A
G
EGFP-POLR1A
KRT20
H
PDO-p18
WT
EGFP-POLR1A
POLR1A High
POLR1A Low
EGFP
FSC-A (x1000)
I
EGFP High
EGFP Low
kDa
EGFP POLR1A
KRT20
ACTIN
J
PDO-p18
POLR1A High
POLR1A Low
Normalized to cell number
EU
K
EU
% of EU+ cells
POLR1A High
POLR1A Low
PDO7
PDO-p18
L
PDO-p18
POLR1A High
POLR1A Low
Normalized to cell number
OP-P
M
OP-P
% of OP-P cells
POLR1A High
POLR1A Low
PDO7
PDO-p18
N
PDO-p18
OP-P POLR1A
KRT20 POLR1A
Str
O
EGFP-POLR1A (PDOs)
Patient derived xenograft (PDX)
Tumor disaggregation
POLR1A high
POLR1A low
PDO-p18
POLR1A High
POLR1A Low
5000 cells
% tumor-free flanks
Time (Days)
PDO7
POLR1A High
POLR1A Low
1000 cells
% tumor free flanks
Time (Days)

Figure 5


PDO-p18
LGR5+ / POLR1A High
LGR5+ / POLR1A Medium
LGR5+ / POLR1A Low
Absorptive-like (KRT20+ LAMC2+)
A
B LGR5High signature
C POLR1AHigh signature
D
Pseudotime
0 10 20 30
E
TSNE 1
TSNE 2
Component 1
Component 2
PDO7
POLR1A High
POLR1A Medium
LGR5+
Enteroendocrine precursors (NEUROG3+)
Enteroendocrine-like (KRT20+ CHRG+)
Mucosecreting-like (MUC2+ KRT20+)
Absorptive-like (KRT20+ LAMC2+)
F
G LGR5High signature
H POLR1AHigh signature
I
Pseudotime
0 10 20 30
J
K
AVVS1 locus
5' HA
Ubq
LoxP
mTagBFP2
STOP
LoxP
tdTomato
3' HA
LGR5/POLR1A/KRT20 locus
ORF
5' HA
LF2A
CreERT2
BGHpA
3' UTR
3' HA
CRISPR Cas9
SubQ injection
50-70 mm3
Tamoxifen
72h
2w
d23/4w
L
Total clones
Number of clones/ tumor area
Day 3
Day 15
Day 23
Day 30
ns
POLR1A
LGR5
KRT20
M
Large clones (>15 cells)
Number of large clones/ tumor area
****
N
Average clone size
Number of cells/clone
POLR1A
LGR5
KRT20
Days
O
Total tomato+ cells
Number of tomato+ cells/ tumor
P
POLR1A-tdTomato
LGR5-tdTomato
d3
d15
d23
d30
Q
KRT20 POLR1A-tdTom
KRT20 LGR5-tdTom

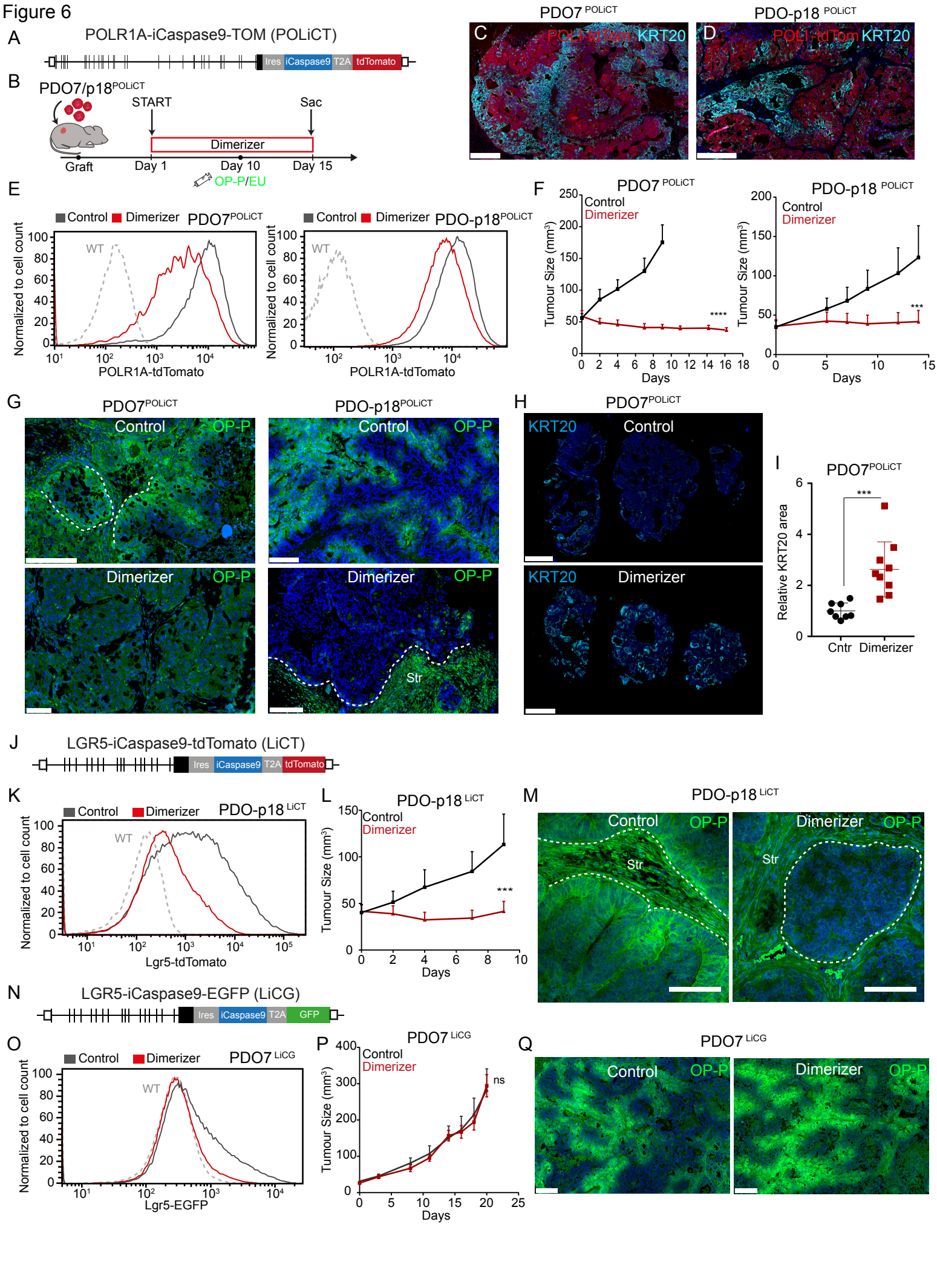


Figure 6
A
POLR1A-iCaspase9-TOM (POLiCT)
Ires
iCaspase9
T2A
tdTomato
B
PDO7/p18POLiCT
START
Sac
Dimerizer
Graft
Day 1
Day 10
Day 15
OP-P/EU
C
PDO7 POLiCT
POLI-tdTom KRT20
D
PDO-p18 POLiCT
POLI-tdTom KRT20
E
Control
Dimerizer
PDO7POLiCT
PDO-p18POLiCT
Normalized to cell count
WT
POLR1A-tdTomato
F
PDO7 POLiCT
PDO-p18 POLiCT
Control
Dimerizer
Tumour Size (mm3)
Days
****
***
G
PDO7POLiCT
PDO-p18POLiCT
Control
OP-P
Dimerizer
Str
H
PDO7POLiCT
KRT20
Control
Dimerizer
I
PDO7POLiCT
Relative KRT20 area
***
Cntr
Dimerizer
J
LGR5-iCaspase9-tdTomato (LiCT)
Ires
iCaspase9
T2A
tdTomato
K
Control
Dimerizer
PDO-p18 LiCT
Normalized to cell count
WT
Lgr5-tdTomato
L
PDO-p18 LiCT
Control
Dimerizer
Tumour Size (mm3)
Days
***
M
PDO-p18 LiCT
Control
OP-P
Dimerizer
Str
N
LGR5-iCaspase9-EGFP (LiCG)
Ires
iCaspase9
T2A
GFP
O
Control
Dimerizer
PDO7 LiCG
Normalized to cell count
WT
Lgr5-EGFP
P
PDO7 LiCG
Control
Dimerizer
Tumour Size (mm3)
Days
ns
Q
PDO7 LiCG
Control
OP-P
Dimerizer

Figure 7

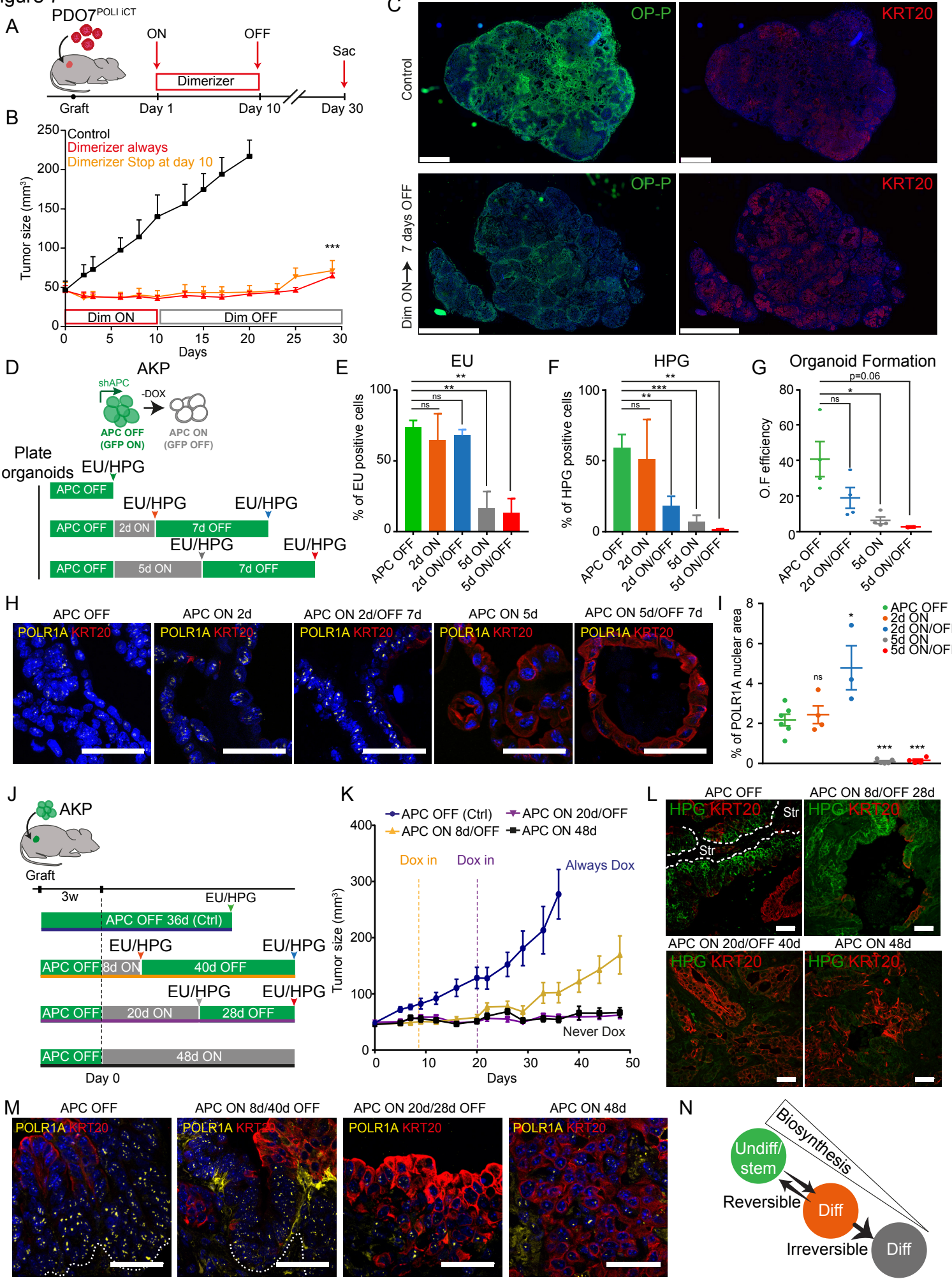

A
PDO7POLI iCT
ON
OFF
Sac
Dimerizer
Graft
Day 1
Day 10
Day 30
B
Control
Dimerizer always
Dimerizer Stop at day 10
Tumor size (mm3)
Dim ON
Dim OFF
Days
***
C
OP-P
KRT20
Control
Dim ON
7 days OFF
D
AKP
shAPC
-DOX
APC OFF (GFP ON)
APC ON (GFP OFF)
Plate organoids
EU/HPG
APC OFF
2d ON
7d OFF
5d ON
E
EU
% of EU positive cells
APC OFF
2d ON
2d ON/OFF
5d ON
5d ON/OFF
F
HPG
% of HPG positive cells
G
Organoid Formation
O.F efficiency
p=0.06
H
APC OFF
APC ON 2d
APC ON 2d/OFF 7d
APC ON 5d
APC ON 5d/OFF 7d
POLR1A KRT20
I
% of POLR1A nuclear area
J
AKP
Graft
3w
EU/HPG
APC OFF 36d (Ctrl)
8d ON
40d OFF
20d ON
28d OFF
48d ON
Day 0
K
APC OFF (Ctrl)
APC ON 8d/OFF
APC ON 20d/OFF
APC ON 48d
Dox in
Dox in
Always Dox
Never Dox
Tumor size (mm3)
Days
L
APC OFF
APC ON 8d/OFF 28d
APC ON 20d/OFF 40d
APC ON 48d
HPG KRT20
Str
M
APC OFF
APC ON 8d/40d OFF
APC ON 20d/28d OFF
APC ON 48d
POLR1A KRT20
N
Biosynthesis
Undiff/ stem
Reversible
Diff
Irreversible
Diff

Figure S1. OP-P and EU Incorporation in CRC models, Related to Figure 1.

(A) OP-P incorporation and KRT20 staining in primary human CRC implanted orthotopically into the caecum of immunodeficient mice. (A') and (A") show representative insets of (A). *Str: Stroma.* Scale bars: in (A): 2.5 mm; in (A'): 250 µm; in (A"): 250 µm.
(B) OP-P incorporation and KRT20 staining in human hepatic metastasis implanted orthotopically into the liver of immunodeficient mice. (B') and (B") show representative insets of (B). *Str: Stroma.* Scale bars: in (B): 2.5 mm; in (B') and (B"): 100 µm.
(C) EU incorporation and KRT20 staining in primary human CRC implanted orthotopically into the caecum of immunodeficient mice. (C') and (C") show representative insets of (C). *Str: Stroma.* Scale bars: in (C): 2.5 mm; in (C'): 100 µm and in (C"): 250 µm.
(D) EU incorporation and KRT20 staining in liver metastasis generated from intrasplenic (IS) injection of PDO7. (D') and (D") show representative insets of (D). *Str: Stroma*. Scale bars: in (D): 2,5 mm in (D') and (D"): 250 µm.
(E) Experimental design for the analysis of OP-P and EU incorporation in EPHB2 tumor cell populations purified from PDXs.
(F) EPHB2 and KRT20 staining (left) and OP-P incorporation (right) in a PDX. Scale bars: 100 µm. Images from (A-D) and (F) are tiled and stitched.

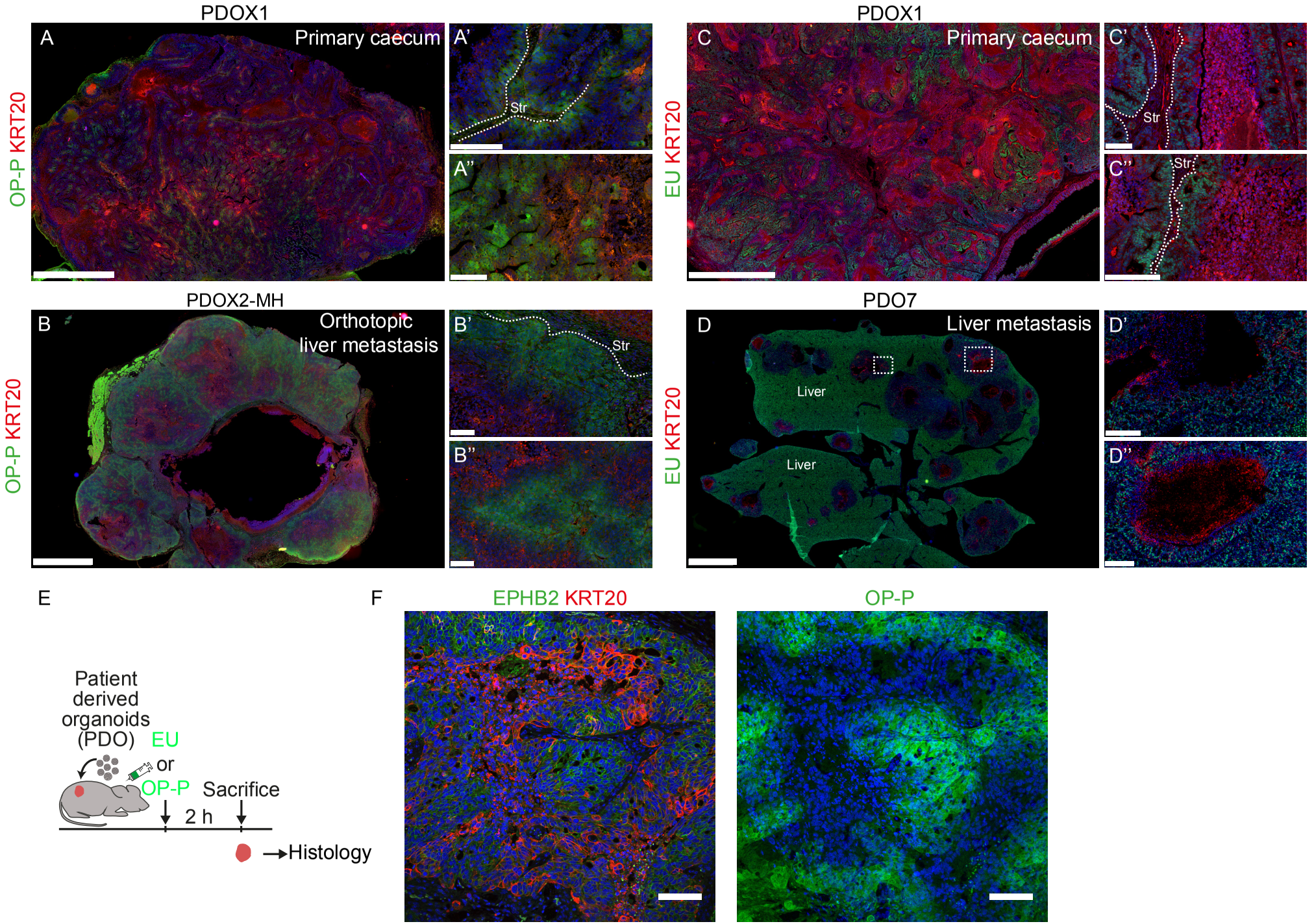

PDOX1
A
Primary caecum
A'
Str
A''
OP-P KRT20
PDOX1
C
Primary caecum
C'
Str
C''
Str
EU KRT20
PDOX2-MH
B
Orthotopic
liver metastasis
B'
Str
B''
OP-P KRT20
PDO7
D
Liver metastasis
Liver
Liver
D'
D''
EU KRT20
E
Patient
derived
organoids
(PDO)
EU
or
OP-P
Sacrifice
2 h
Histology
F
EPHB2 KRT20
OP-P

Figure S2. Proliferation is not Limited to the Tumor Periphery in CRC, Related to Figure 2.

(A and **B)** Ki67 staining in surgically resected human CRC tumors implanted into the caecum of immunodeficient mice. (A') and (B') show in more detail the center of the tumor, and (A") and (B") show the periphery. Scale bars in (A) and (B): 2.5 mm; in (A'), (A"), (B'), and (B"): 250 μm.

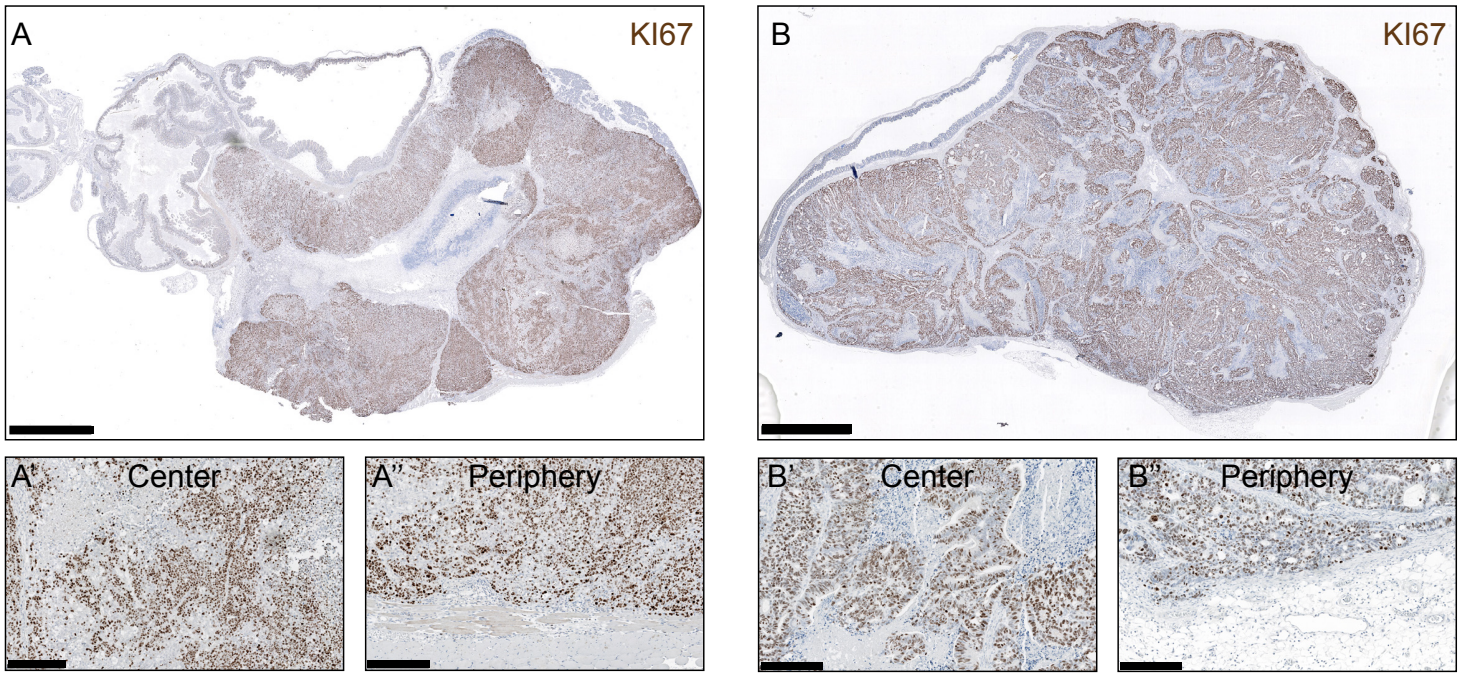
(C) Experimental design for the quantification of KI67 positive tumor cells in different areas of the tumor. The panel indicates selected areas for quantification: three layers of the periphery (P1-P3) and a central region (Center).

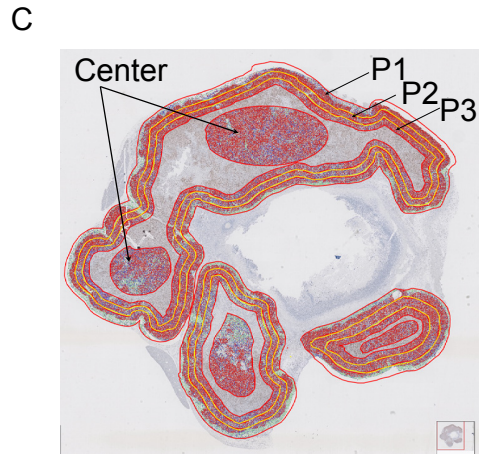
(D) Quantification of Ki67 positive tumor cells in the indicated areas of 5 different orthotopic tumors (T2-T6).

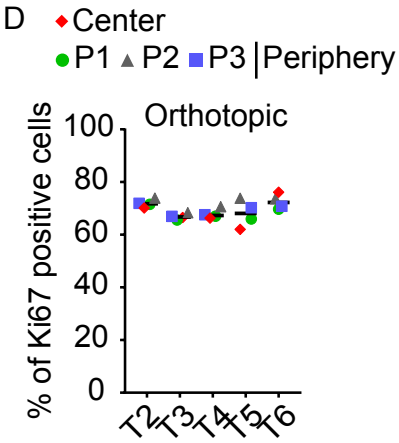
(E) Representative Ki67 staining in primary human CRC. (E') shows in more detail the center of the tumor and (E") shows the periphery. Scale bars in (E): 2.5 mm; in (E') and (E"): 250 μm.

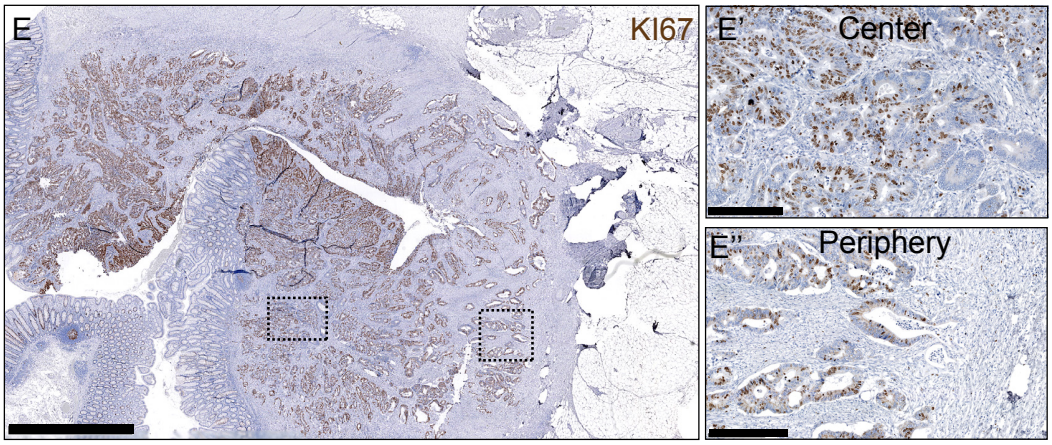
(F) Quantification of Ki67 positive tumor cells in 4 different primary human CRC samples. The same strategy as in (C) was applied.

Figure S2-Related to Figure 2









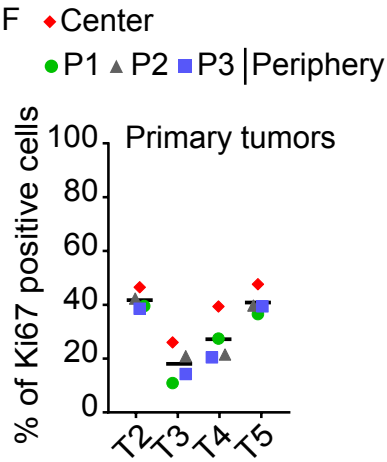


Figure S3. Downregulation of rDNA Transcription and Loss of Ribosomal Content upon Tumor Cell Differentiation, Related to Figure 3.

(A and **B)** EU incorporation and KRT20 expression in SW403 tumor cells (A) and in AKP organoids (B) upon *in vitro* differentiation at indicated time points. Scale bars: 50 μm.
(C) 5.8S and KRT20 staining of 6 representative primary human CRC samples. Scale bars: 1mm.
(D) Quantification of KRT20 and 5.8S (Y10B antibody) intensity of 24 analyzed primary human CRC tumors. For the analysis, each sample was split into 10 μm^2 tiles and the intensity of KRT20 and 5.8S was quantified and represented in dot plots. The figure shows the resulting plots of each quantified tumor.

Figure S3-Related to Figure 3

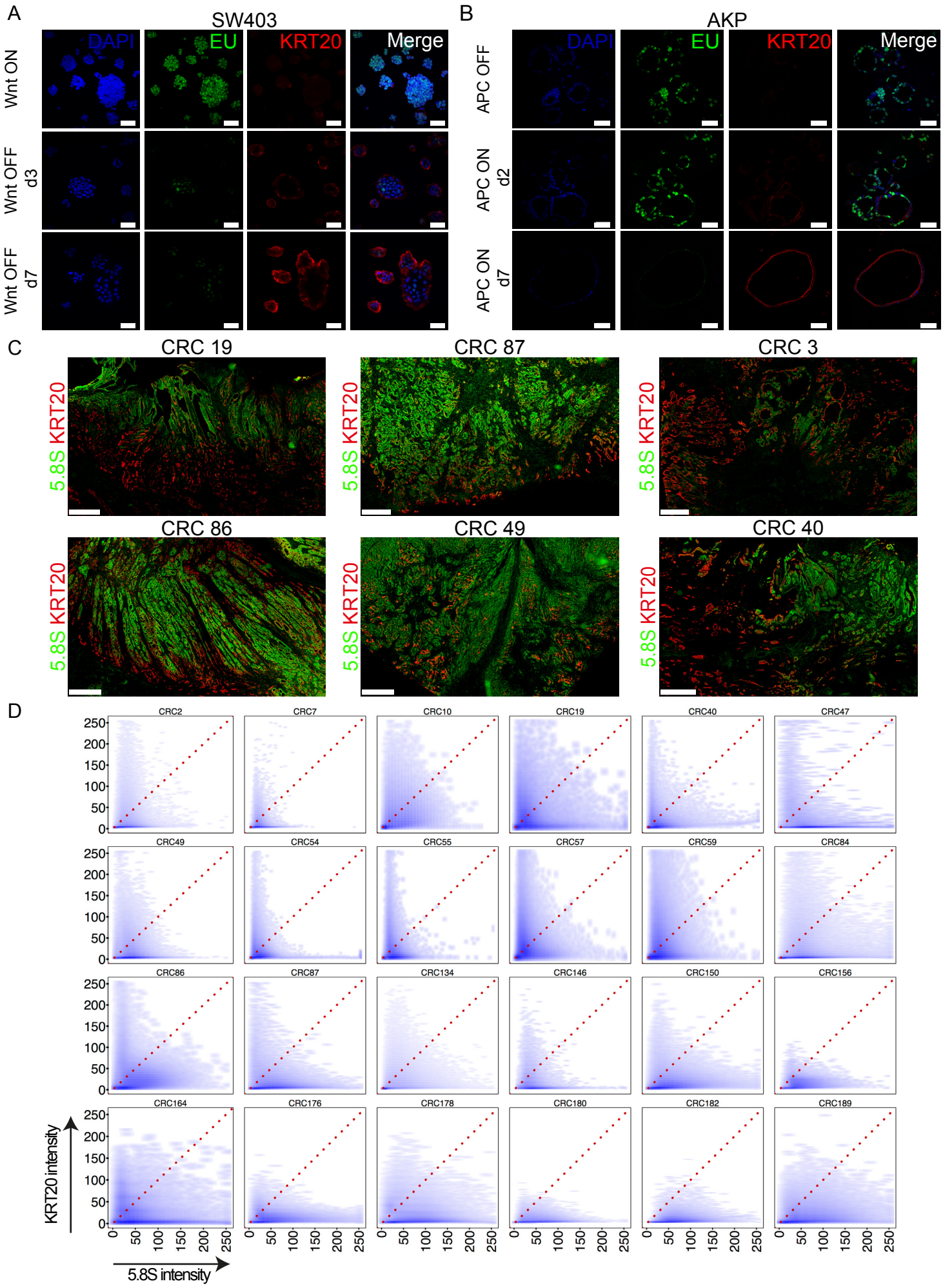


Figure S4. Characterization of POLR1A Tumor Cell Population *in vitro* and *in vivo*, Related to Figure 4.

(A) POLR1A and KRT20 staining in control (Wnt ON) or differentiated (Wnt OFFF) CRC SW403 cell line. Scale bars: 20µm. Images are tiled and stitched.
(B and **C)** POLR1A and KRT20 staining in tumor-derived xenografts from PDO7 (B) and PDO-p18 (C). White squares demark representative insets. Scale bars: 100 µm. Images are tiled and stitched.
(D) POLR1A gene expression in human colon normal mucosa (n=50) compared to colorectal adenocarcinomas (n=98). Association between POL1RA expression and sample type was evaluated using a linear model (Wald test).
(E) *In vitro* colocalization of POLR1A protein in red (left) and endogenous EGFP-POLR1A (right) in knock-in organoid PDO7. Scale bars: 20µm.
(F) EGFP-POLR1A and KRT20 staining in tumor-derived xenografts from injected knock-in PDO7. Scale bars: 50 µm.
(G) Representative flow cytometry plot of EGFP-POLR1A expression in untargeted (left) or in knock-in PDO7 tumor xenograft (right). Squares indicate the gating used for sorting EGFP-POLR1A populations.
(H) RT-qPCR analysis of *EGFP* and *POLR1A* in sorted EGFP-POLR1A population indicated in (G). Bars depict the mean and the upper/lower limits of relative expression from a representative experiment.
(I) Representative flow cytometry plots of EU (left) and OP-P (right) incorporation in EGFP-POLR1A populations sorted from tumor-derived xenografts from injected knock-in PDO7.
(J) Organoid formation efficiency of EGFP-POLR1A populations sorted from tumor-derived xenografts from injected knock-in PDO7 (left) and PDO-p18 (right). **p=0.0095, ****p<0.0001 in an unpaired two-tailed t-test. Confidence intervals are mean±SEM. Right panels show representative images of organoid formation efficiency of POLR1A population in PDO-p18. Scale bar: 100 µm.
(K) Experimental set up for the metastasis initiation capacity experiment in (L).
(L) Quantification of liver photon flux (photons per second) after intraspleen injection of 45.000 sorted cells of POLR1A-tdTomato high and low tumor cells purified from tumor-derived xenografts from injected knock-in PDO7. Right panel: representative images of liver metastases at the end point. White arrowheads indicate metastases.

(M) Knock-in constructs for the generation of double EGFP-POLR1A and LGR5-iCaspase9-tdTomato (LiCT) knock-in PDO7.

(N) Percentage of POLR1A high and low cells within LGR5+ and LGR5- tumor cell populations from double knock-in PDO7.

(O) Gene set enrichment analysis (GSEA) of LGR5+ versus LGR5- within POLR1A high tumor cells. For all hallmarks p-val<0.05.

(P) Gene set enrichment analysis (GSEA) of the mTORC1 (left) and Wnt signaling (right) pathways in LGR5+ POLR1A high versus LGR5- POLR1A high tumor cells.

Figure S4-Related to Figure 4

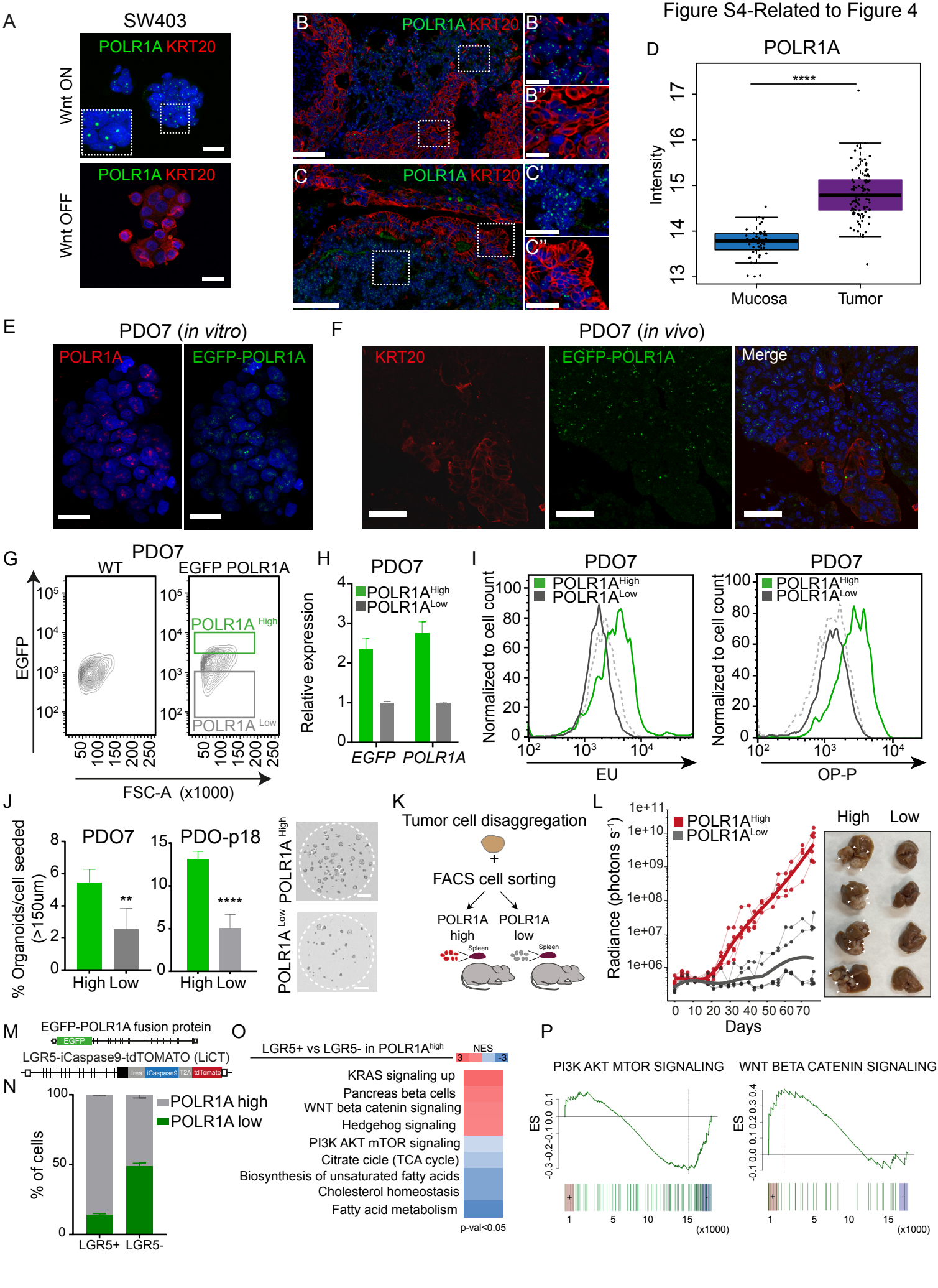


Figure S5. scRNA-seq and lineage tracing analysis of POLR1A, LGR5 and KRT20 tumor cell populations, Related to Figure 5.

(A and **B)** Cluster-specific plots of 388 (PDO-p18)(A) and 511 (PDO7)(B) single cells from tumor xenograft displaying the relative expression of ISC markers (LGR5, SMOC2), POLR1A low signature and differentiation markers of absorptive (LAMC2), enteroendocrine (KRT20, NEUROG3) and mucosecreting (MUC2) lineages.

(C and **E)** POLR1A high signature color-coded by the pseudo-time ordering of cells of PDO-p18 (C) and PDO7 (E).

(D and **F)** LGR5 signature, LGR5 mRNA, POLR1A high and POLR1A low gene signature across the different cluster identities in PDO-p18 (D) and PDO7 (F).

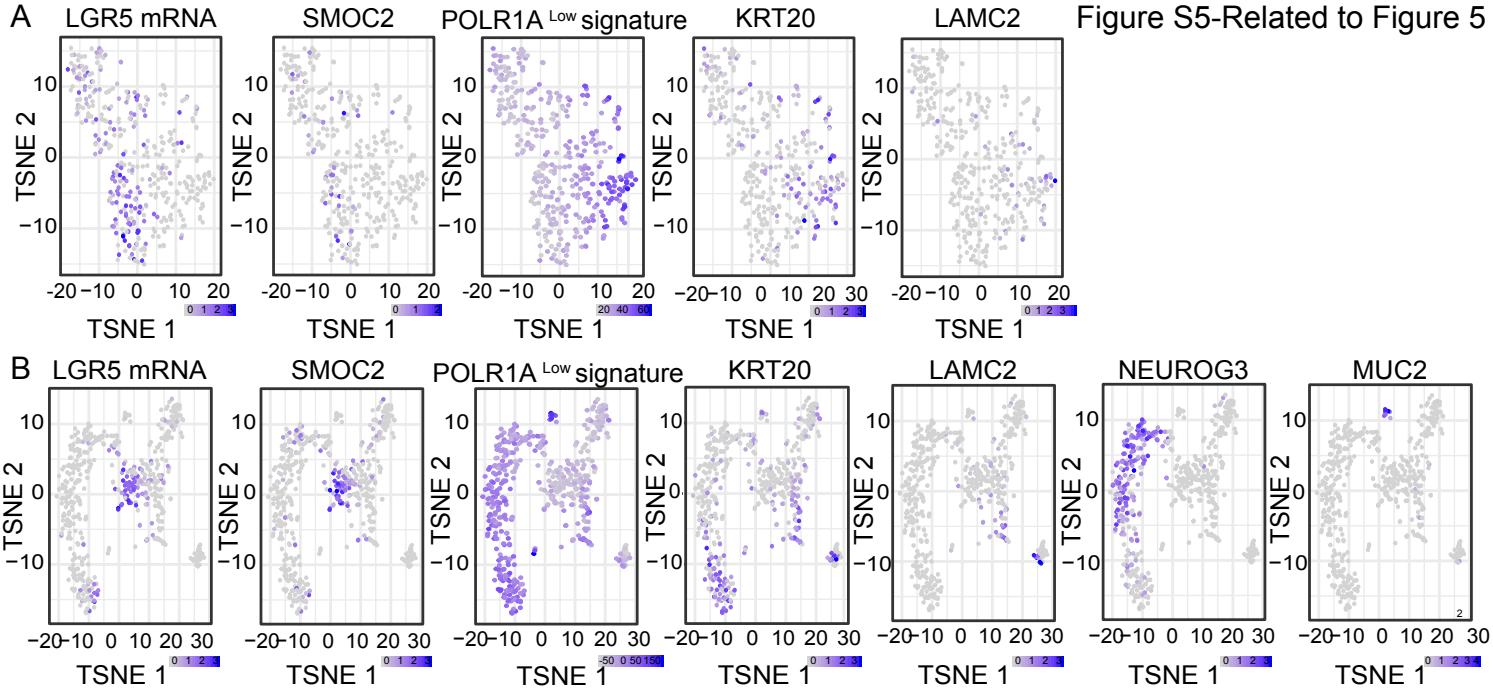
(G) Representative flow cytometry plots of tdTomato positive cells in tumor xenorafts 3 days after tamoxifen induction in the three different populations. Right panel represents the average of tdTomato positive cells for the different tumors analyzed in each population (POLR1A: n=3, LGR5: n=4, KRT20: n=5). $p>0.05$ in an unpaired Tukey's multiple comparison test. Confidence intervals are mean±SEM.

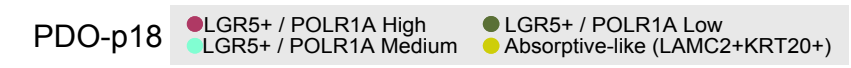
(H) Average ratio of viable epithelium over time for all tumors analyzed in the tracing experiments for each population.

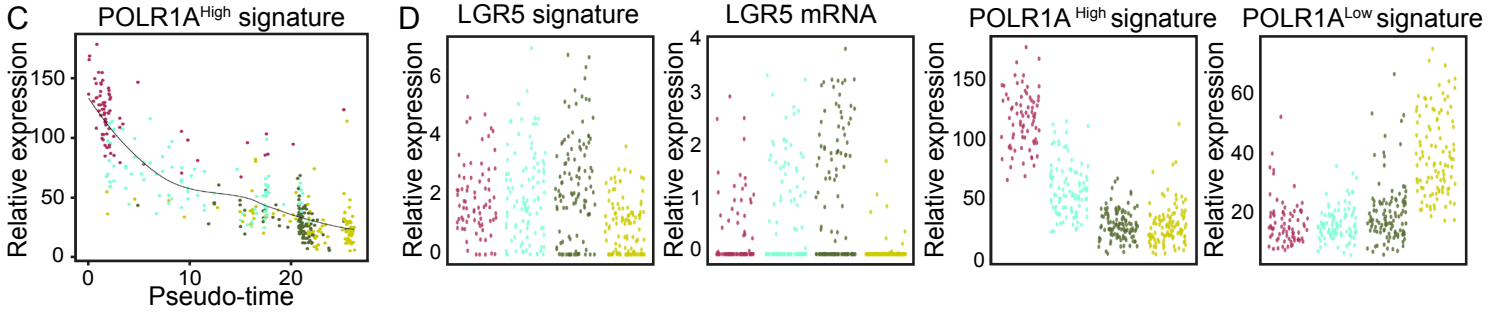
(I) KRT20 and tdTomato staining in tumor sections 3 days after tamoxifen induction of KRT20 tracing. Arrowheads show single tdTomato positive cells (left) or double positive tdTomato/KRT20 positive cells (right). Images are tiled and stitched.

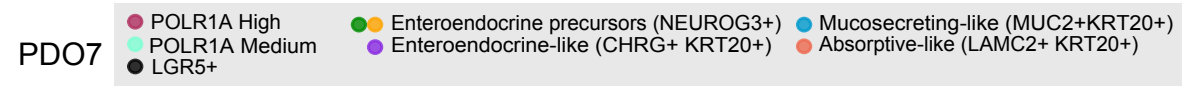
(J) Percentage of double (tdTomato and KRT20) and single (tdTomato) positive cells at day 3 post tamoxifen induction. (T1-T4 are individual tumor xenograft).

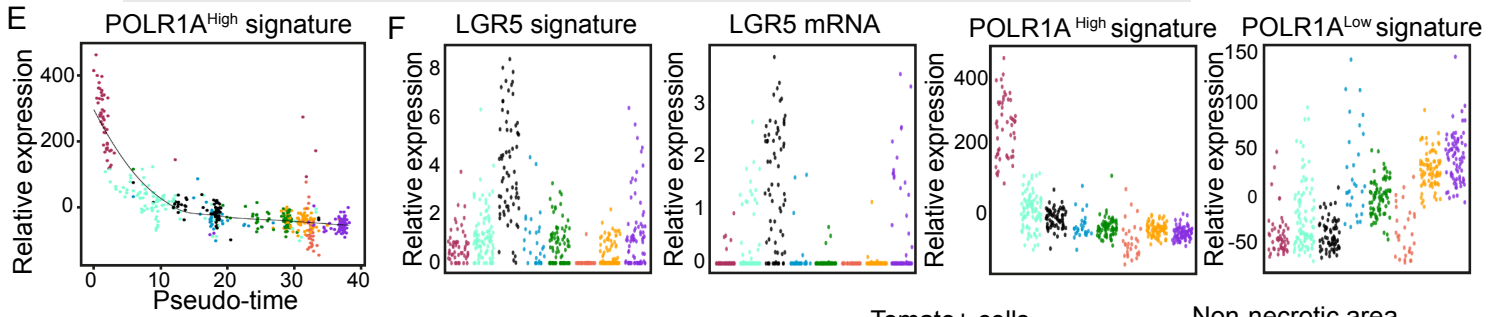
Figure S5-Related to Figure 5

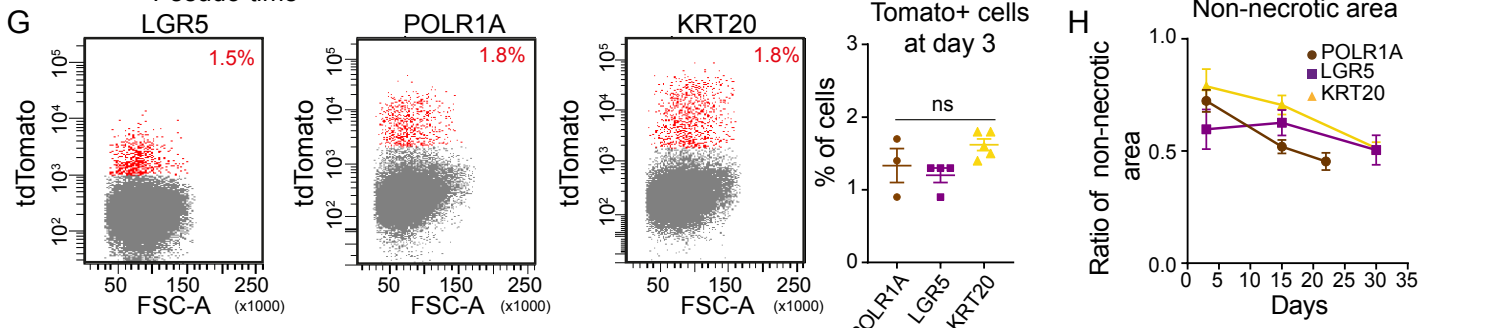


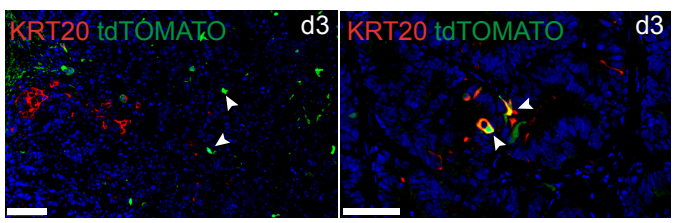












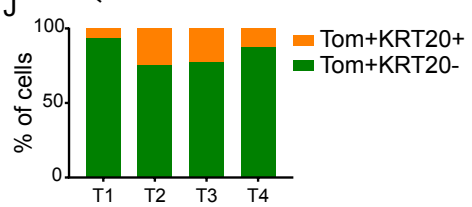


Figure S6. Ablation of POLR1A tumor cells induces tumor growth arrest and cell differentiation, Related to Figure 6.

(A) RT-qPCR analysis of *tdTomato* and *POLR1A* in POLR1A-iCT (POLiCT) high and low populations purified from *in vivo* knock-in PDO7 and PDO-p18. Bars depict the mean and upper/lower limits of relative expression obtained from a representative experiment.
(B) Representative images of tdTomato staining in tumor-derived xenografts from injected POLiCT knock-in PDO7 (left) and PDO-p18 (right) under control conditions (upper panels) or after 5 days of POLR1A ablation (lower panels). tdTomato is shown in rainbow. Scale bars: 2.5 mm, 500 µm and 1 mm.
(C) Quantification of OPP (left panel) and EU (right panel) positive cells in POL-iCT high and low populations purified from *in vivo* knock-in PDO7. For OP-P *p=0.015 (n=2 mice) and for EU **p=0.0018 (n=5 mice) in a paired two-tailed t-test. Confidence intervals are mean±SEM.
(D) EU incorporation in untreated tumors (control) or after 10 days of POLR1A tumor cell ablation (dimerizer) in PDO7 and PDO-p18. Scale bars: 250 µm.
(E) Representative images of PAS and Alcian blue staining in control or after 10 days of POLR1A ablation in PDO7. Scale bars: 2.5mm
(F) Quantification of Alcian blue positive area. For PDO7: ***p=0.0004 (Controls n=8 and Dimerizer n=7 tumors). For PDO-p18: p=0.054 (Control n=4 and Dimerizer n=5 tumors) in an unpaired two-tailed t-test. Confidence intervals are mean±SEM.
(G and **H)** Representative images showing the efficiency of LGR5 ablation in PDO7 knocked-in with LiCG (G) or PDO-p18 knocked-in with LiCT (H). Scale bars: 2.5 mm.
(I) Experimental design of the chemotherapy treatment.
(J) Absolute tumor growth during chemotherapy treatment (Folfiri) in PDO-p18 (left panel) and PDO7 (right panel). Animals bearing subcutaneous xenografts were injected intraperitoneally with a single dose of Folfiri at indicated time points. At endpoints for PDO-p18 ***p=0.0002 (Controls n=4 and Folfiri n= 8 tumors). For PDO7 ****p<0.0001 (Controls n=7 and Folfiri n=8 tumors) in an unpaired two-tailed t-test. Confidence intervals are mean±SEM.
(K) Representative flow cytometry plots showing tdTomato expression in knock-in POLiCT PDO-p18 (upper panel) and PDO7 (lower panel) at the end of chemotherapy treatment. Dashed line represents unlabelled parental tumors (WT).

(L) OPP incorporation and KRT20 expression in PDO-p18 untreated tumors or after 10 days of Folfiri treatment. Scale bars: 250 μm. Images are tiled and stitched.

(M) Quantification of the relative KRT20 area in control and Folfiri treated tumors at end point in (J). For PDO7: *p=0.01 (Controls n=6 and Folfiri n=9 tumors). For PDO-p18: ***p=0.0002 (Controls n=7 and Folfiri n=7 tumors) in an unpaired two-tailed t-test. Confidence intervals are mean±SEM.

(N) Western blot analysis of POLR1A, KRT20 and SDCBP2 levels in PDO7 upon 24h of *in vitro* BMH-21(1μM) treatment.

(O) RT-qPCR analysis of ISC and differentiation genes in PDO7 after 24h of *in vitro* BMH-21(1μM) treatment. Bars depict the mean and upper/lower limits of relative expression obtained from a representative experiment.

(P) KRT20 staining of PDO7 organoids after 24h of *in vitro* BMH-21 treatment. Scale bars: 100μm. Images are tiled and stitched.

(Q) Absolute tumor growth during BMH-21 treatment. Animals bearing subcutaneous xenografts were injected intaperitoneally daily with 100 mg/Kg of BMH-21. At end points *p=0.026 (Control=8 and BMH21=8 tumors) in an unpaired two-tailed t-test. Confidence intervals are mean±SEM.

(R) Quantification of relative KRT20 area in control and BMH-21 treated tumors at the end point in (Q). p=0.06 (Controls n=8 and BMH21 n=8 tumors) in an unpaired two-tailed t-test. Confidence intervals are mean±SEM.

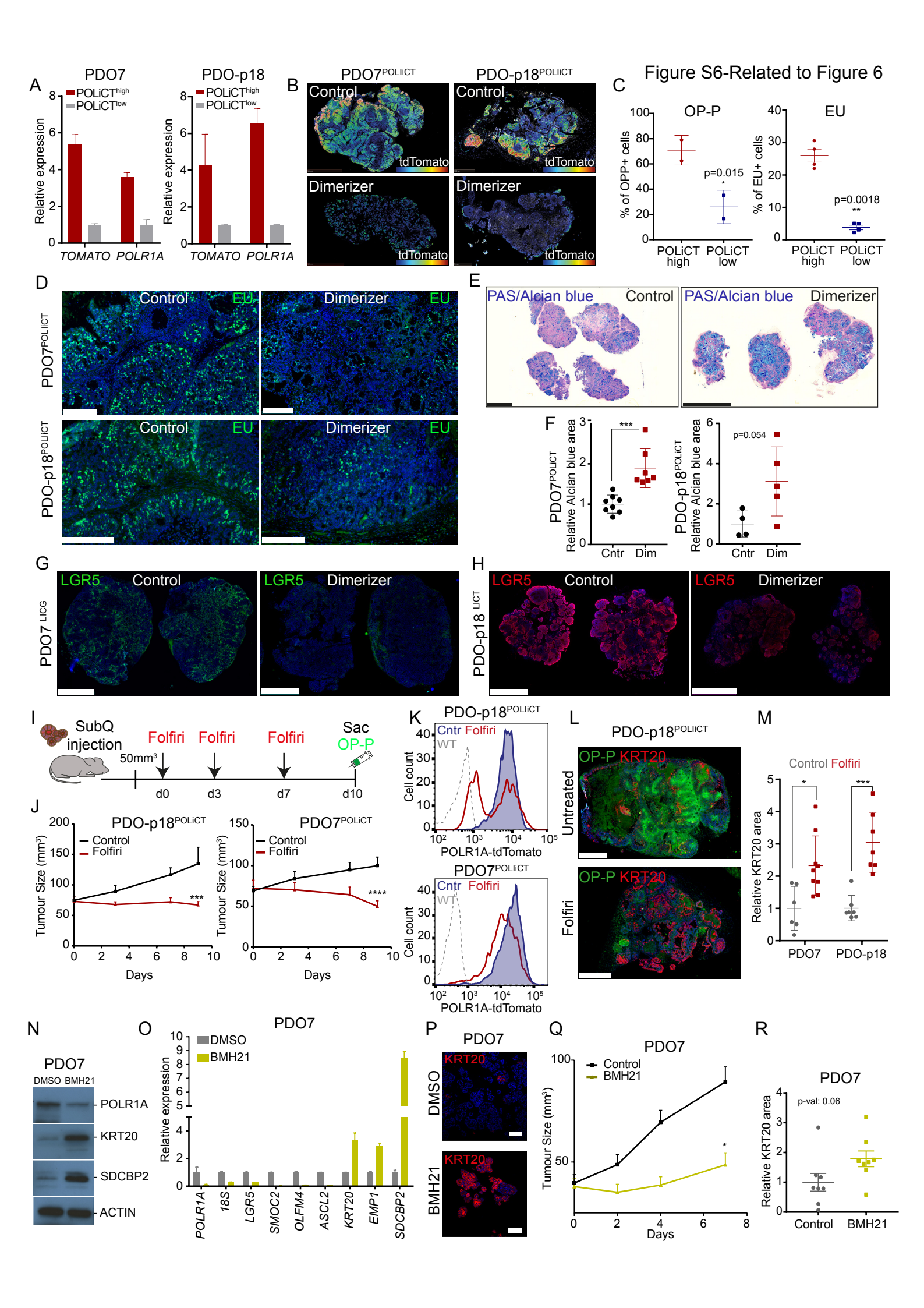
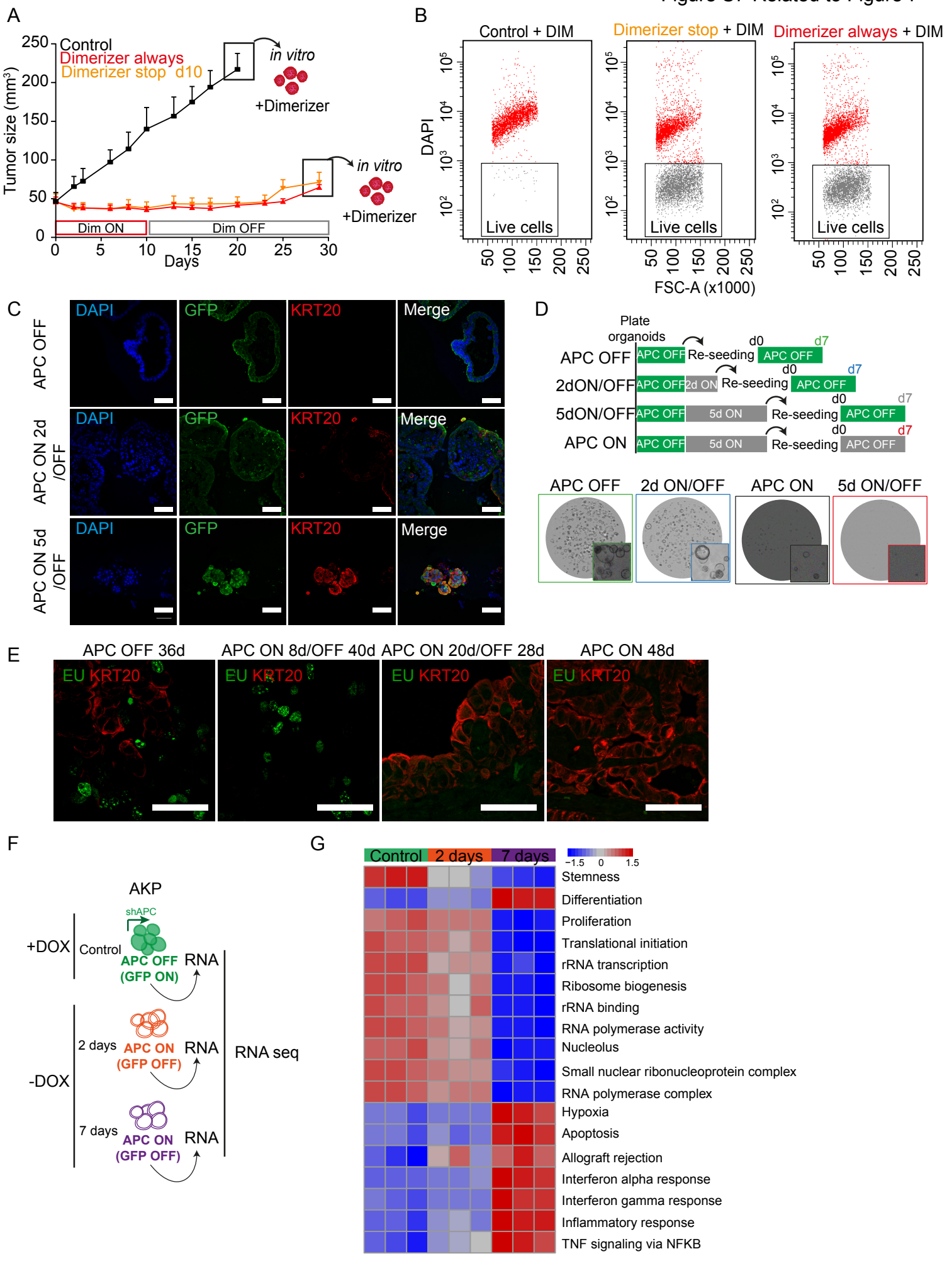


Figure S6-Related to Figure 6
A
PDO7
PDO-p18
POLiCT high
POLiCT low
Relative expression
TOMATO
POLR1A
B
PDO7 POLiCT
PDO-p18 POLiCT
Control
Dimerizer
tdTomato
C
OP-P
EU
% of OPP+ cells
% of EU+ cells
p=0.015
p=0.0018
POLiCT high
POLiCT low
D
PDO7 POLiCT
PDO-p18 POLiCT
Control
Dimerizer
EU
E
PAS/Alcian blue
Control
Dimerizer
F
PDO7 POLiCT
PDO-p18 POLiCT
Relative Alcian blue area
***
p=0.054
Cntr
Dim
G
PDO7 LICG
LGR5
Control
Dimerizer
H
PDO-p18 LICT
I
SubQ injection
50mm3
Folfiri
Folfiri
Folfiri
Sac
OP-P
d0
d3
d7
d10
J
PDO-p18 POLiCT
PDO7 POLiCT
Control
Folfiri
Tumour Size (mm3)
Days
***
****
K
PDO-p18 POLiCT
PDO7 POLiCT
Cntr
Folfiri
WT
Cell count
POLR1A-tdTomato
L
PDO-p18 POLiCT
Untreated
Folfiri
OP-P KRT20
M
Control Folfiri
Relative KRT20 area
PDO7
PDO-p18
N
PDO7
DMSO
BMH21
POLR1A
KRT20
SDCBP2
ACTIN
O
PDO7
DMSO
BMH21
Relative expression
POLR1A
18S
LGR5
SMOC2
OLFM4
ASCL2
KRT20
EMP1
SDCBP2
P
PDO7
KRT20
DMSO
BMH21
Q
PDO7
Control
BMH21
Tumour Size (mm3)
Days
R
PDO7
p-val: 0.06
Relative KRT20 area
Control
BMH21

Figure S7. Irreversible Tumor Cell Differentiation, Related to Figure 7

(A) Tumors that grew under dimerizer treatment *in vivo* were dissociated and cells were plated *in vitro* to test the acquisition of dimerizer resistance.
(B) Viability flow cytometry plots of control or resistant cells (dimerizer stop and dimerizer always) treated with dimerizer *in vitro*.
(C) GFP (shAPC reporter) and KRT20 expression in AKP organoids that were differentiated (APC ON) during 2 or 5 days and recovered for 7 more days (APC OFF). Notice that after 5 days of APC ON reversion is not possible and organoids remain differentiated. Scale bars: 50 μm.
(D) Experimental design of the organoid formation experiment. Representative images of AKP organoid formation efficiency experiments during *in vitro* differentiation (ON) and after reversion (ON/OFF).
(E) EU was injected into animals bearing subcutaneous AKP organoids that were either 8 or 20 days without doxycycline (APC ON) and were recovered for an additional 40 or 28 days (APC OFF). Images show EU incorporation in control (APC OFF 48d), reverted (APC ON 8 or 20/OFF) and differentiated (APC ON 48d) tumors. Scale bars: 50 μm. Images are tiled and stitched.
(F) Experimental design of (G).
(G) Gene set enrichment analysis (GSEA) in AKP organoids in control (shAPC OFF) versus early (2 days) or late (7 days) differentiation (shAPC ON).

Figure S7-Related to Figure 7



Supplementary Tables 1-4

Table S1. sgRNA guides, Related to STAR Methods.

Locus-insertion	sgRNA sequence
POLR1A ATG	CTTGGAGATCAACATCCTCC
POLR1A 3'UTR	GAGCTGGGCAGATGGTGCCG
LGR5 3'UTR	TGTCTCTAATTAATATGTGA
KRT20 3'UTR	ATATCTAAATAGCTACCAGA
AAVS1-BFPTOM	GTCCCTAGTGGCTAAGGC

Table S2. Genotyping primers, Related to STAR Methods.

Locus-insertion	5' specific	3' specific
EGFP-POLR1A	F: CTGATAACTTCTTTAGCCAAAGT R: GCTGAACTTGTGGCCGTTTA	F: ACATGGTCCTGCTGGAGTTC R: AGATGTACAGACAGCTGAACAAAT
POLR1A-iCT	F: AAAGCTCCTGATTGATAACTTGG R: AAGACGGCAATATGGTGGAAAAT	F: GCATTGTCTGAGTAGGTGTCATT R: AATACATTTCCCTGGCATGGACA
LGR5-iCG/T	F: GTTTTTTGTCCATTTTTGCTTCAG R: GCTACAGACGTTGTTTGTCTTCAA	F: ATGCTGGGGATGCGGTGGGCTCTAT R: AGTGGAATTCCCTTCTGAGCTTTG
POLR1A-LF2A-CreERT2	F: AAAGCTCCTGATTGATAACTTGG R: CGAACATCTTCAGGTTCTGCGGG	F: TTCCTTGCAAAAGTATTACATCAC R: AATACATTTCCCTGGCATGGACA
LGR5-LF2A-CreERT2	F: GTTTTTTGTCCATTTTTGCTTCAG R: CGAACATCTTCAGGTTCTGCGGG	F: TTCCTTGCAAAAGTATTACATCAC R: AGTGGAATTCCCTTCTGAGCTTTG
KRT20-LF2A-CreERT2	F: GCACCCTGGAAGAGAGAGGTAAG R: CGAACATCTTCAGGTTCTGCGGG	F: TTCCTTGCAAAAGTATTACATCAC R: CCGTAGAGACAGGAAGTAGATTAG
AAVS1-BFP-TOM	F: GGACCACTTTGAGCTCTACTGG R: GGGCCATTTACCGTAAGTTATGTA	F: GCATTGTCTGAGTAGGTGTCATTC R: ATGAGATGGTGGACGAGGAAGG

Table S3. Southern blot probes, Related to STAR Methods.

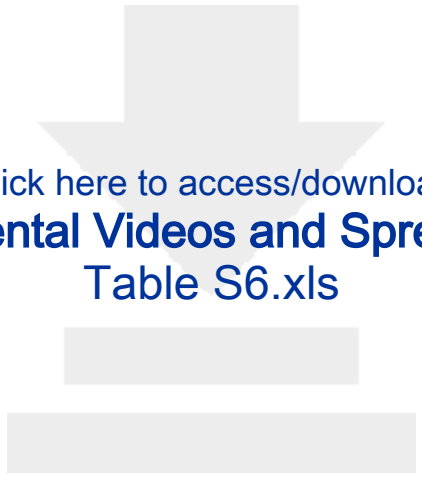
Reporter	Primer	Restriction enzyme
EGFP	F: CCACCATGGTGAGCAAGGGCGAGG R: TTACTTGTACAGCTCGTCCATGCC	BglII
TdTOMATO	F: GGGCATGGCACCGGCAGCACC R: CCTACTTGTACAGCTCGTCCATGCC	KpnI

Table S4. PDO genotypes, Related to Figures 1-6

PDO	WNT pathway	KRAS pathway	TGF-B pathway	PI3K pathway	TP53/ ATM
PDO7	APC STOP (R787*)/STOP (K1438*)	KRAS (G13/+)	SMAD4 (L536R /L536R)	WT	ATM (V182L / N1983S)
PDO-p18	APC (deletion/splice-site)	WT	WT	WT	WT (insensitive to Nutlin-3a)*
PDO19b	RNF43 (A355fs)	BRAF (V600E)	WT	PIK3CA (H1047R)	TP53 (Q331_splice) (R273C)
PDOX1	APC p.D1483fs	KRAS p.A146V	n.d.	n.d.	TP53 p.R282W
PDOX2	No information available				

Table S5

Table S6



Click here to access/download

Supplemental Videos and Spreadsheets

Table S6.xls

Table S7